Universidad Autónoma de Madrid

Tesis Doctoral

Effects of group size on social decisions and the role of Histone acetylation in behavioral individuality, two views on animal collectives

Autor:
Julián Vicente Page

Director:
Dr. Gonzalo García de Polavieja Embid

Memoria de Tesis doctoral presentada
para la obtención del título de
Doctor
por la Universidad Autónoma de Madrid
Programa de Doctorado en Biofísica
Instituto Nicolás Cabrera
Facultad de Ciencias

Junio de 2015

A mis padres

Contents

List of Figures

Summary

In this thesis I have addressed two different topics related to animal collectives: group decision making and behavioral variability. Group decisions have been observed in many animal species and it has been shown that animal collectives use the information of other conspecifics to improve the accuracy of their decisions. Behavioral variability occurs in animal populations even in absence of genetic or environmental differences, and its underlying molecular mechanisms remain elusive.

In Chapter 1, we present and analyze a new model of collective decision making. This model allows individuals to interchange information by reevaluating their choices in a decision making task. The opportunity of a change of opinion increases the accuracy of the decision by correcting the mistakes of the first deciding individuals, which have less access to social information. The model predicts the existence of an optimal group size in dichotomous decision making tasks. The presence of an optimal group size has been observed in different animal species and has been theoretically postulated by ecological models in terms of individual fitness. In our case, the existence of the optimal group size is independent of the mechanisms or the formula that individual use to make the decision.

Our model is also the first model that, having correlation between the votes of the individuals, improves the accuracy of Condorcet's Theorem. Moreover, it is able to explain a set of parameters of a decision making experiment in fish, confirming its validity for understanding real animal data. We also conducted different spatio-temporal simulations to prove that the existence of an optimal group size is inherent to any decision making process that includes the possibility of a change of opinion.

In Chapter 2, we use an experimental model in larval zebrafish to study the molecular mechanisms that modulate behavioral individuality. We show that the

behavior of larvae is stable over days by analyzing two different parameters: the total time of movement (activity) and the mean distance to the center of the well in which they swim (radial index). The intra-individual variability of these parameters is lower than the inter-individual variability, so our setup can be used to study variability among the individuals of a population.

We show that the use of sodium butirate or anacardic acid, inhibitors of the activity of Histone Deacetylase and Histone Aacetyltransferase respectively, reduce the behavioral variability in a larval population. This directly implies Histone acetylation in the generation of behavioral variability in zebrafish. By ChIP sequencing, we identified the genomic regions whose hypervariable acetylation state was related with behavioral variability. These regions defined a gene network implied in several neurodevelopmental processes and its variability, measured as the coefficient of variation of mRNA expression across the population, correlated with the behavioral variability.

We also found that transcription factor Yin-Yang 1 was present in these regions and that its binding correlated with the amount of behavioral variability of the population. We confirmed the importance of this transcription factor by showing that a Yin-Yang 1 heterozygotic mutant also presented a reduced molecular and behavioral variability. Finally, we show that the levels of acetyl coenzyme A , an important molecule in metabolism, also correlate with the behavioral variability of the population.

Resumen

En esta tesis he tratado dos temas diferentes relacionados con colectivos de animales: la toma de decisiones en grupo y la variabilidad comportamental entre los individuos de una población. Las decisiones en grupo son un fenómeno que se ha observado en muchas especies animales, en las que además se ha probado que los colectivos de animales utilizan información de otros miembros del grupo para mejorar el resultado de la decisión. La variabilidad en el comportamiento ocurre en poblaciones de animales incluso en ausencia de diferencias genéticas o de medio ambiente, y los mecanismos moleculares implicados aún se conocen con poco detalle.

En el Capítulo 1, presentamos los resultados de un nuevo modelo de toma de decisiones colectivas. Este modelo permite el intercambio de información entre los individuos ofreciéndoles la posibilidad reevaluar sus elecciones en una tarea de toma de decisiones. La oportunidad de cambiar de opinión aumenta la precisión de la decisión gracias a la corrección de los posibles errores de los individuos que decidieron con menos información social. El modelo predice la existencia de un tamaño óptimo de grupo en tareas de toma de decisión entre dos opciones. La presencia de tamaños óptimos de grupo se ha observado en diferentes especies animales y ha sido postulada por modelos teóricos en términos de fitness de los individuos. En nuestro caso, la existencia de un tamaño óptimo no depende del mecanismo específico o de la fórmula que los individuos usen para tomar la decisión.

Nuestro modelo es además el primero que, aun teniendo una correlación entre los votos de los individuos, mejora el resultado del Teorema de Condorcet. Más aún, es capaz de explicar una serie de parámetros de un experimento de toma de decisiones en peces, confirmando su validez a la hora de estudiar el comportamiento grupal en experimentos de animales. Además, realizamos una serie de simulaciones espaciotemporales para probar que la existencia del tamaño óptimo es inherente

a cualquier proceso de toma de decisiones en grupo que incluya la posibilidad de un cambio de opinión.

En el Capítulo 2, utilizo un modelo experimental en larva de pez cebra para estudiar los mecanismos moleculares que modulan la individualidad comportamental. Mostramos que el comportamiento de las larvas es estable a lo largo de los días analizando dos parámetros diferentes: el tiempo total que las larvas están en movimiento (actividad) y la distancia media al centro del pocillo en el que nadan (índice radial). La variabilidad entre individuos de la población es menor que la intra-varriablidad del propio individuo, con lo que nuestro modelo experimental se puede utilizar para estudiar la variabilidad comportamental entre individuos de una misma población.

También mostramos que el butirato sódico o el ácido anacárdico, inhibidores de la actividad de la Histona Deacetilasa y de la Histona Acetiltransferasa respectivamente, reducen la variabilidad comportamental en una población de larva de pez cebra. Este resultado involucra directamente la participación de la acetilación de histonas en la generación de variabilidad comportamental. Mediante el uso de secuenciación masiva, identificamos las regiones cuya hipervarialibidad en acetilación está relacionada con la variabilidad en el comportamiento. Estas regiones definen una red génica implicada en procesos de neurodesarrollo y su variabilidad, medida mediante el coeficiente de variación de la expresión en mRNA a lo largo de la población, correlaciona con la variabilidad en el comportamiento.

Describimos, además, cómo el factor de transcripción Yin-Yang 1 está presente en estas regiones y que su unión correlaciona con la cantidad de variabilidad comportamental de la población. Confirmamos su importancia mostrando que un mutante heterocigótico de Yin-Yang 1 también presenta una variabilidad molecular y comportamental reducida. Para concluir, muestro que los niveles de acetil concima A, una molécula muy importante en el metabolismo, también correlacionan con la variabilidad comportamental de la población.

1. General introduction

Animal societies are a widespread phenomenon and can be found in all sort of species. The type of society changes dramatically from one species to the other and so does their internal organization. Eusocial insects like ants, honeybees or termites, show an extreme form of sociality, involving highly organized societies, with individual organisms specialized for distinct roles (Nowak et al., 2010; Thorne, 1997; Weinstock et al., 2006). Other species organize themselves in different kinds of societies and display a variable repertory of social interaction, such as communal living, cooperative care of young, or division of labour. The range of species included in this group is extremely wide and goes from avian societies (Cockburn, 1998) to mammals like meerkats (Brotherton et al., 2001) or lions (Packer et al., 1990). This thesis deals with two different topics that can only be understood in the context of animal collectives: group decision making and behavioral individuality. The aim of this introduction is to introduce the reader to some particular aspects of animal groups that will help to put this thesis in context.

There is a wide range of advantages that individuals can get from joining others, and we can find examples in all kinds of animal societies. In some species, for example, breeding individuals rely on the assistance of nonbreeding helpers to raise the young, even though they may not be directly related (Clutton-Brock, 2002). Individuals can also use information from other members of the group to exploit evenly distributed food supplies more effectively (Galef and Wigmore, 1983; Ward and Zahavi, 1973) or they can use conspecifics to maximize their own daily feeding rate (see Kruuk (1972) for cooperative hunting situations or Davies and Houston (1981) for territory defense economics).

Animals also experience a decrease in predation risks when living in group, due to several reasons. For instance, it has been shown that predator attacks become less successful (Kenward, 1978) or that vigilance increases in different

animal species (Bertram, 1980; da Silva and Terhune, 1988; Santema and Clutton-Brock, 2013). A more protective mechanism is social defense, which has also been shown to decrease the performance of predator attacks (Bode et al., 2010; Haas, 1985). Another advantage that reduces the probability of being attacked by a predator is the dilution effect, which just states the fact that the larger the group of prey animal is, the smaller the chance that any particular individual being the victim (Dehn, 1990; Foster and Treherne, 1981). Dilution effect is a direct consequence of the dynamics of the individuals and their mutual interaction when avoiding a predator.

A very different spatial interaction between individuals arises when they forage together. It has been shown that group foraging improves the efficiency of the foraging, either by saving energy (Weimerskirch et al., 2001) or by improving its accuracy (Biro et al., 2006; Greene, 1987). As an example of a foraging task I show three frames of a video, recorded at the laboratory, of a group of medaka fish (*Oryzias latipes*) exploring a novel setup (Pérez-Escudero et al., 2014).

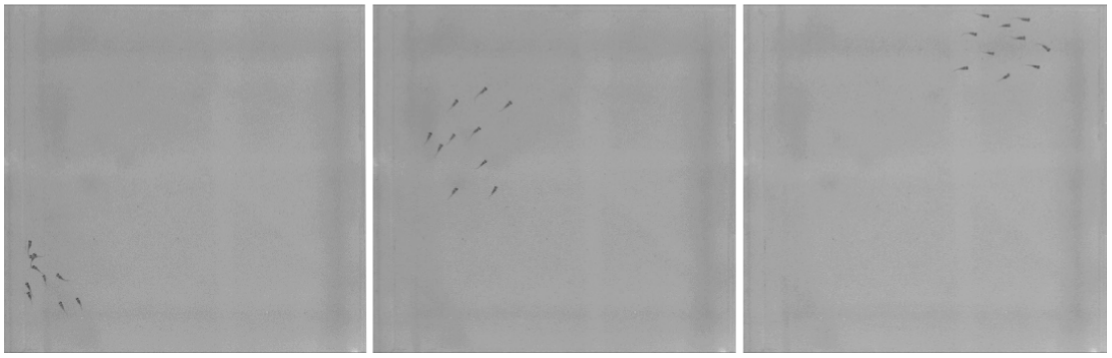

FIGURE 1.1: Three non-consecutive frames of a five second sequence of 10 medaka (Oryzias latipes) exploring a novel setup. The video lasted for 10 minutes and fish performed the exploration in a coordinated and high directional way, called schooling.

Group foraging is a clear example of collective decision making. Groups of animals often need to make communal decisions in a variety of activities, and their choices rely on the decision of others. Nest choice of ants or bees is a well studied example, where the decision mechanism involves a number of more informed individuals, or 'scouts', communicating their 'opinions' to the rest of the colony (Seeley, 1997). Eventually, a consensus decision is reached by all group members after weighting and pooling the available information of all the scouts (Branco et al., 2006; Franks et al., 2003; Seeley and Buhrman, 1999). Other examples of collective decisions arise when groups of primates have to decide where to travel

after a rest period (Stewart, K.J. and Harcourt, 1994) or when to initiate a specific activity (Boinski and Campbell, 1995; Milton, 2000). Social decision making will be the central topic of Chapter 2, where I will talk about its relation with an optimal size in animal groups.

From an evolutionary point of view, there is still an ongoing debate about the reasons why animals live in group (Pievani, 2014). Natural selection acting only at the level of the individual is not able to explain the origin of cooperative or altruist behaviors, and many mechanisms have been proposed to make cooperation possible (Nowak, 2006). Examples of these mechanisms are kin selection (an evolutionary strategy that favours the reproductive success of an organism's relatives), different types of reciprocity (where individuals may benefit from mutual interaction), group selection (which states that natural selection acts at the level of the group) or multilevel selection (that stands in favour of selection acting at several levels, from genes to groups). Although these mechanisms offer different explanations, they all agree that, in many situations, group living increases the fitness of the whole group while decreasing the fitness of the individual.

This decrease in the individual fitness comes from the fact that living in group has also some obvious costs associated. The limitation of resources required for survival may generate a competition among animals of the same species, known as intra-specific competition. The scarcity of food caused by weather changes or other environmental reasons creates a intra-specific competition in different animal species, from invertebrate to mammals (White, 2008). Sexual selection, acting on males or females of the same species, has been shown to be determinant for the individual rate of reproduction (Clutton-brock, 2007). Other less obvious costs, such as increased parasite have also been demonstrated (Brown and Brown, 1986). The costs of group living have not been studied empirically as much as the benefits, in part because it is reasonable to assume that as group size increases, eventually so does competition for resources (Sumpter, 2010).

It has recently been argued that these negative effects of group living can be counterbalanced by the phenotypic diversity of the individuals in a population (Bolnick et al., 2011; Hughes et al., 2008). Ecologist increasingly recognize individual variation as an important factor affecting intra-specific competition (Bolnick et al., 2011, 2003; Hughes et al., 2008; Wolf and Weissing, 2012). Traits such as boldness, aggressiveness or activity are often directly related to mortality risk

and/or fecundity (Biro and Stamps, 2008; Réale et al., 2010; Smith and Blumstein, 2008; Wolf and Weissing, 2012), so differences in this features can result in distinct individual fitness values. Furthermore, phenotypic variation can also compensate possible environmental changes, and species become more resilient and less vulnerable (McCann, 2000; Oldroyd and Fewell, 2007; Schindler et al., 2010).

Variability can also regulate intra-specific competition by dispersal of the individuals in a population (Duckworth, 2008; Saastamoinen et al., 2009). In this sense, individuals with special characteristics may be able to face some specific environmental conditions that other individuals may not. Thus, variation in different traits affects not only the movements between habitats, but also the distribution of individuals within habitats, leading to an spatial segregation expected to reduce exploitation competition (Duckworth, 2006; Wolf and Weissing, 2012). Behavioral variability will be addressed in Chapter 3, where I will show our results about an epigenetic pathway related to the generation of behavioral variability.

The nexus between the two chapters of my thesis is animal behavior. Behavior is often defined as the activity or movement of an organism (Plotkin, 1988; Tinbergen, 1963) and is a key factor mediating the interactions of individuals with their environment (Duckworth, 2009; Wcislo, 1989). These interactions determine where organisms live and how they obtain resources, avoid predators, choose mates and respond to conspecific and heterospecific competitors. Behavior is then an important output that helps in understanding the immediate response of an individual to its environment. In this Thesis I studied behavior from two different but connected branches. Neither group behavior nor variability can be understood outside animal collectives and both aspects are intimately related and they probably regulate each other.

The approach I have used to address these questions has been completely interdisciplinary and, in my opinion, this is one of the strongest points of this Thesis. Interdisciplinarity is needed for the study of animal behavior, as behavior connects effects acting at several levels, including genes, neural structures and physiological responses. The study of behavior is then a field where questions of physics, biology, psychology and social sciences converge (Gomez-Marin et al., 2014). In this sense, there are different examples of the tools that scientist use when studying behavior, from genetic manipulation to high technical devices. The latter ones have improved substantially in recent years, giving researchers the access to new

technologies such as high resolution videos, accelerometers or GPS. One of these techniques is a new tracking software that I helped developing in the lab (see (Pérez-Escudero et al., 2014) and that is currently been used by several laboratories in their studies. From a different point of view, the study of animal behavior has been also approached by the use of mathematical modeling, presenting very useful results (see, for example, Arganda et al. (2012); Couzin et al. (2005a) and Ward et al. (2008a)). Finally, the classical use of pharmacological treatments or genetic manipulation is essential for understanding the molecular mechanisms underlying behavior (Nelson and Chiavegatto, 2001; Robinson et al., 2008). These tools are just examples of the ones I have used in this Thesis, and they will be discussed through the following chapters.

Personally, my background in theoretical physics helped me through the mathematical modeling, data analysis and problem solving. All my knowledge about molecular biology comes from the years I spent developing this research and, even though my background is still weak and I have thousands of things to learn, I consider myself capable of supporting the essence of all the biological results of this Thesis. The main general results, which I proceed to present, will, in my humble opinion, be appealing to different fields of the scientific community.

2. Optimal group size in collective decision making

2.1 Introduction

As I referred to in the General Introduction, living in groups not only benefits their members in various ways, but also results in costs that will negatively affect the fitness of the individuals of the group (Krause and Ruxton, 2002; Sumpter, 2010). One example of these costs is the competition for resources that increases dramatically when the number of individuals within the group rises. Therefore, the expected costs of living in group will be higher than the benefits for groups of large sizes. On the contrary, animals in small groups, where costs of living together are reduced, may not be able to benefit of the advantages of group living (e.g. social defense or cooperative breeding). Thus, by weighing both the costs and the benefits of living in group, the fitness function of each individual is assumed to have a maximum at an intermediate group size, generally called the optimal group size. It is assumed to be different for each species because their specific ecological conditions will define different cost and benefit functions (Couzin and Krause, 2003; Krause and Ruxton, 2002).

A theoretical formalization of the existence of an optimal group size based on the fitness function is given in (Brown, 1982). In this study, the author modeled the fitness function assuming that, in territorial animals, the costs of being in a group came from resource depletion and defense expenses while the benefits were obtained by simply sharing these defense expenses. Another classical approach was used in Sibly (1983), Giraldeau and Gillis (1985) or Higashi and Yamamura (1993), raising some interesting questions about the stability of the optimal group size and the existence of different group sizes in nature. These authors used simple models of aggregation where an individual animal would decide to join a group as

long as its fitness inside the group was higher than that of being alone. In spite of their importance for applying mathematical models to the existence of an optimal group size, their strong assumptions about the form of the fitness function and the lack of experimental data make their predictions difficult to test (Martinez and Marschall, 1999).

Some experimental studies have shown the existence of an optimal group size by studying the lifetime reproductive success in a wide range of species. Female social spiders (*Anelosimus eximius*) raised a maximum number of offspring in intermediate colony sizes (Avilés and Tufiño, 1998) and group sizes of lions (*Panthera leo*) were found to typically remain within the range that maximized individual reproductive success (VanderWaal et al., 2009). Another work on northern bobwhite, *Colinus virginianus*, showed a reduction on the individual survival for large or small coveys (Williams et al., 2003) and Brown (1996) found that cliff swallows (*Petrochelidon pyrrhonota*) in colonies of between 30 and 80 nests produced more surviving young than smaller or larger colonies. It has also been shown that stress levels, measured as fecal cortisol concentration, are lower in groups of lemurs (*Lemur catta*) of intermediate size and that the frequencies of roe deer (*Capreolus capreolus*) group sizes had a maximum at 3 individuals, independently of the population density (Gerard et al., 2002).

Other studies show how individuals can adjust the size of the group depending on the environmental factors they perceive (see Hoare et al. (2004) and Zöttl et al. (2013)). The first authors claim that individuals can assess the size of the group by making individual decisions based on local interactions with others. This mechanism, in which individual decisions create an outcome for the complete group, has more generally been studied from the point of view of collective decision making. The core of this chapter relies on the confluence between the existence of an optimal group size and the fact that different animal species in nature make communal decisions.

There have been several approaches to the study of collective decision making. The first one was performed by Marquis de Condorcet who, in the 18^{th} century, applied the theory to the problem of improving the jury system (Condorcet, 1785). He found out that, for binary choices in which each individual has a probability p of making a correct decision, the probability that the majority of the group chooses correctly increases steeply with group size. This result is known as Condorcet's Theorem and it holds as long as p is greater than 0.5, provided that individuals vote

independently. The restriction of independent voting has been widely studied and it has been shown that a correlation between the votes of the individuals always lowers the accuracy of Condorcet's Theorem (Berg, 1993; Ladha, 1992; Nitzan and Paroush, 1984).

The findings of this theorem are unlikely to explain how collective decisions are taken in animals, as the process required for independent voting is too artificial to be found in nature. However, they have created a positive starting point for the development of other kinds of models, in which the probability of choosing the correct option increased with the number of individuals committed to that option. The most relevant works in this line have been conducted in groups of fish and they combine both quantitative data and mathematical modeling of two choice experiments (Sumpter et al., 2008, 2011; Ward et al., 2008b). The models show quorum rules that allow fish to make more accurate decisions as group size increases, in accordance with what has been experimentally observed. All of them have the same intrinsic structure, in which each individual chooses one of the options with a certain probability that depends, upon other parameters, on the number of individuals that have chosen that option. They all work sequentially, meaning that individuals decide in order, so the first deciding individual faces the decision on its own, with no information from the rest of the group. This is a limiting factor of these kinds of models and I will go through it in this chapter. Apart from that, none of these models present a formal derivation that could help to understand the meaning of its parameters, and this point was addressed in our laboratory, by developing a collective decision making model directly from Bayesian estimation theory (Arganda et al., 2012; Pérez-Escudero and de Polavieja, 2011).

In this chapter I am using the decision rule from Pérez-Escudero and de Polavieja (2011) to create a completely general decision making model that predicts the existence of an optimal group size. I will compare its results with that of Condorcet's Theorem and I will also show that our model presents a further insight on a set of experimental data that the existing models were unable to explain completely.

2.2 Results

2.2.1 Change of opinion improves the accuracy of a collective decision

We are going to study the paradigm of a group of individuals confronted with a dichotomous choice where one option is better than the other. We created a decision making model in which individuals do not make the decision based on one single estimation about the quality of the options, but instead they can vote several times for their preferred option. This allows individuals to change their opinion and take into account the votes of the rest of the group as the process advances. We will consider that the final decision of each individual will correspond to their last vote in the sequence. For simplicity, we chose the formula from Pérez-Escudero and de Polavieja (2011) as the decision rule, because it clearly integrates the contribution of the social and the private information when computing the probability of voting for a specific option. Each individual votes for option x or option y depending on its own estimation and on the number of individuals that have previously chosen each of the two options, N_x and N_y,

$$P_y = \left(1 + a \cdot s^{-(N_y - N_x)}\right)^{-1} \tag{2.1}$$

with a and s parameters representing the influence of private and social information, respectively. The values of these parameters are going to depend entirely on the circumstances of the decision process, and they integrate information such as the ease of the decision (in parameter a) or the confidence towards other individuals (in parameter s).

For clarity, we first considered the very simple case in which every individual votes twice in a sequential order, keeping this same order in both sequences. In other words, after the last individual in the first sequence has voted, there is a second sequence of votes, that keeps the same order, in which each individual reconsiders its choices following the same decision rule (equation 2.1). In figure 2.1A we show an example of 8 individuals where only six voted for the correct option in the first round but, with the influence of the opinion of the rest of the group, all of them end up choosing the correct option after the second round. It is remarkable that each individual computes the probability of going to each

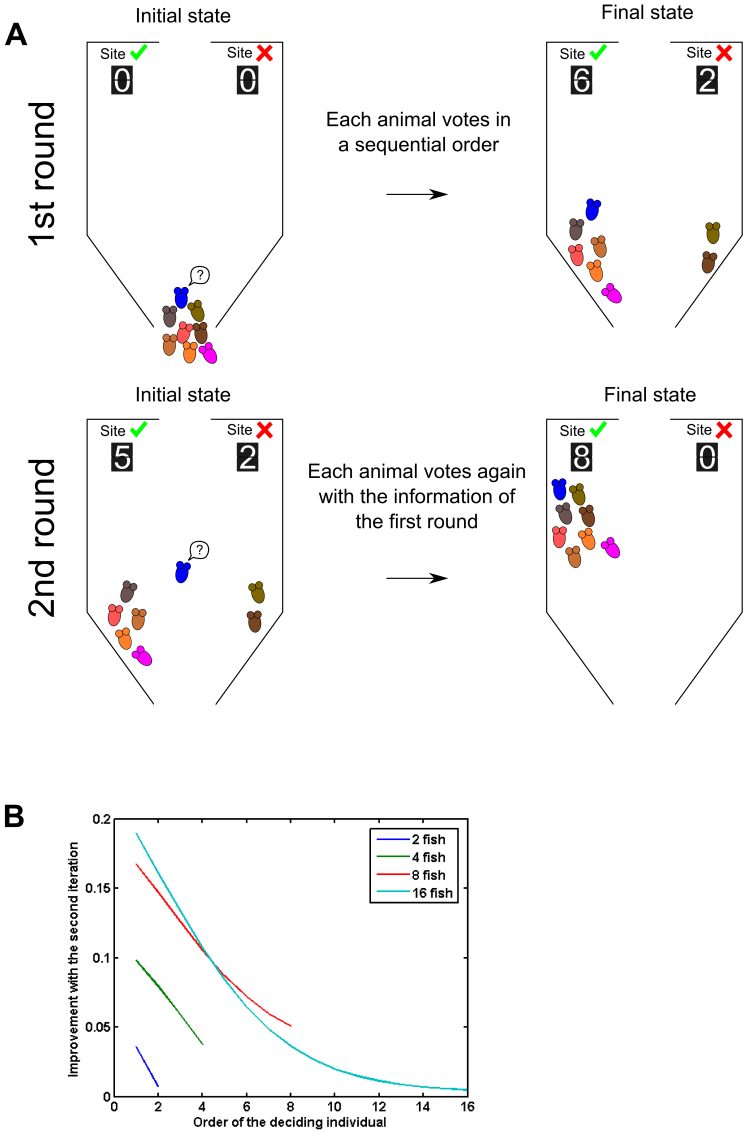

FIGURE 2.1: (A) An example of our decision making model. In the first round each animal decides sequentially computing the probability of going to each option depending on the decision of previous deciding animals. (B) Improvement on the decision for each fish, calculated as the difference between the probability of choosing the correct option at the first and at the second round. The parameters we have used for this plot are $a = 0.45$ and $s = 1.65$.

option with formula 2.1, so the first voting individual is the only one that has no access to social information, as this social information increases while the decision progresses.

For clarity, we first considered the very simple case in which every individual votes twice in a sequential order, keeping this same order in both sequences. In other words, after the last individual in the first sequence has voted, there is a second sequence of votes, that keeps the same order, in which each individual reconsiders its choices. In figure 2.1A we show an example of 8 individuals where only six voted for the correct option in the first round but, with the influence of the opinion of the rest of the group, all of them end up choosing the correct option after the second round. It is remarkable that each individual computes the probability of going to each option with formula 2.1, so the first voting individual is the only one that has no access to social information, as this social information increases while the decision progresses.

Effectively, individuals at the beginning of the first round have little information from the choices of others but our model allows them to have access to this information in the second round. The individual probability of choosing the best option should then be expected to change from the first to the second sequence. We then computed, for each individual, the difference between the probability of choosing the best option at the second and at the first sequence, for 4 group sizes (see Figure 2.1B). We found that the improvement on the estimation of the first deciding individual is higher for larger groups. However, the individual improvement decreases with the order in the sequence and, for larger groups, the last deciding individuals improves less in larger than in smaller groups. In fact, from the fifth individual, the improvement for the group of 16 is lower than for the group of 8, and this improvement still decreases until the last individual.

This effect seems to be caused by the fact that individuals in large groups have to wait for the votes of the rest of the individuals before reevaluating their initial opinion. If this was the case, the decisions of the first individuals should be more determinant for large groups than for small groups. We then decided to compare the group performance between a group of 8 and 16 individuals, provided that the first two individuals choose either the correct or the wrong option (see Figure 2.2). In Figure 2.2A we plotted the probability of voting for the correct option given that the first two individuals have voted for the wrong one. In accordance with our explanation, we found that it is much easier to revert the vote of the two initial

individuals for a group of 8 than for a group of 16. However, the amplification when the first two individuals vote for the correct option is not much higher for the group of 16 (see 2.2A). This shows that error amplification of initial mistakes is much higher for bigger groups but also that they do not experiment the same improvement when amplifying correct initial decisions. These results illustrate not only that our model with changes of opinion can improve the performance of a decision making process, but also that this improvement is higher for intermediate groups, suggesting the existence an optimal group size.

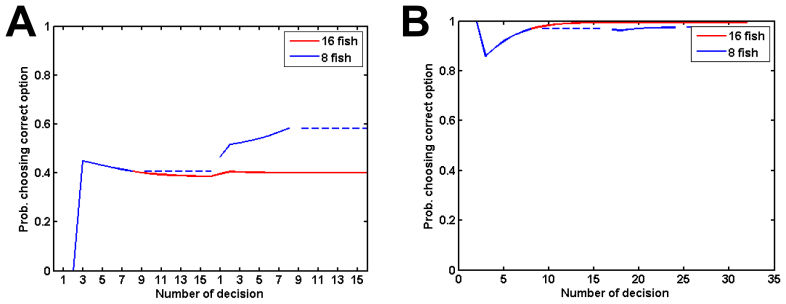

FIGURE 2.2: Probability of choosing the correct option for a group size of 8 and 16 individuals when the first two have chosen either the wrong option (A) or the correct one (B). We see that the difference in the improvement between the group of 8 and the group of 16 when the first two individuals choose the wrong option is smaller than the difference between the group of 16 and the group of 8 when the same animals choose the correct option. This implies that the capacity of overcoming initial mistakes of an intermediate group is higher than the amplification of initial correct choices at large group sizes. The parameters used are the same as in Figure 2.1B.

2.2.2 Prediction of the existence of an optimal group size

If we compute the proportion of individuals choosing the correct option for several voting rounds, we find the existence of an optimal group size for any number of rounds greater than one, see Figure 2.3.A. In addition, the proportion of individuals choosing the correct option increases with the number of rounds, amplifying the presence of the optimum. To gain an intuition on the origin of its optimum, we studied how the probability of each possible final configuration of choices evolves with the number of rounds. In other words, we analyzed how the probability that a certain proportion of individuals chooses the correct option changed with the number of rounds (see Figure 2.3B). For a group of 4 individuals, we found that the probabilities soon become stabilized and that there are several surviving configurations where at least one of the individuals chooses the wrong option. This does

not happen for the group of 8, as the probability of every individual choosing the correct option increases at each round, reducing the probability of finding other configurations with less correct choices. However, for the group of 16, we found that the probabilities become stabilized again but now with only two surviving configurations: one with all the individuals in the correct option and the other with all of them in the wrong option. Both probabilities of these states increase from the first round to the others, meaning that such a big group size amplifies what the majority has done in the initial estimation sequence, as we saw in 2.2.

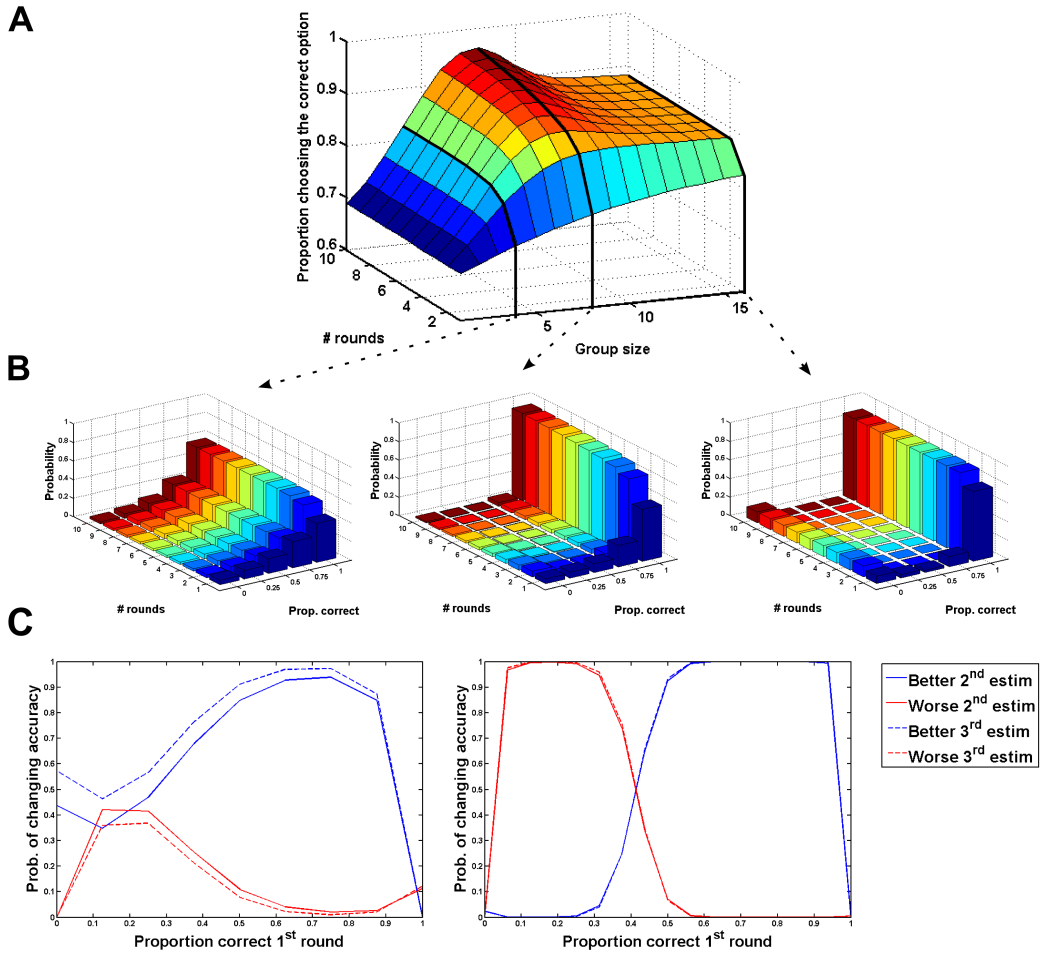

FIGURE 2.3: (A) Proportion of individuals choosing the correct option, plotted for several group sizes and for 10 voting rounds. The parameters used are the same as in Figure 2.1B, $a = 0.45$ and $s = 1.65$. (B) Evolution of the different outcomes of the experiment through different voting rounds, for group sizes of 4, 8 and 16. (C) Capacity of correcting/amplifying the results of the first round, for group sizes of 8 (left) and 16 individuals (right).

To prove this argument in a more general way, we studied the probability of increasing the performance depending on what individuals have done on the first

sequence. More explicitly, for every initial configuration, we compute the probability of having a better performance (Figure 2.3C, blue lines) and that of having a worse performance (Figure 2.3C, red lines) in the second round. For the group of 8 individuals, the probability of having a better performance in the second round is in general higher than that of having a worse performance, independently of the results of the first one. In fact, this effect gets amplified in the third round, where the probability of choosing the correct option increases for every initial case (Figure 2.3C, blue dashed lines). However, this does not happen for the case of 16 individuals, where we see two collapsed states with a very low probability of moving from one to the other. If a high proportion of individuals chose wrong in the first sequence, it would make the whole group chose the wrong option, without any possibility of correcting the initial mistake. The same happens if the majority of individuals go to the correct option, as the whole group would finally end choosing it. Therefore, we find that a big group has such a social cohesion that it is going to amplify what individuals have done in the first round, while an intermediate group has still some chance of correcting mistakes and, at the same time, gets the advantage of having enough social information.

2.2.3 The existence of an optimum is inherent to several round voting

We wanted to discard that the presence of the optimum was caused by the exact parameter combination of our formula, so we performed a scan through these parameters and we found not only that the optimum does not disappear, but also that it changes its value for each parameter combination (see Figure 2.4). If the weight of the social information is high (high values of s) then the effect of the private information of each individual becomes irrelevant and, consequently, the value of the optimal group size is generally low. However, if the individuals become less influenced by the decisions of others (low values of s), the probability of amplifying initial errors become less definitive, so the optimal group size tends to increase. Although the effect of the private information is less influential, we find that for individuals with low private information (high values of a), the probability of initial errors is considerably high, and in this case is more beneficial to avoid error amplification and to be in smaller group sizes. However, when the private information becomes considerably high (low values of a), the interaction with

others can be beneficial, as we just saw in section 2.2.2, so the optimal group size increases.

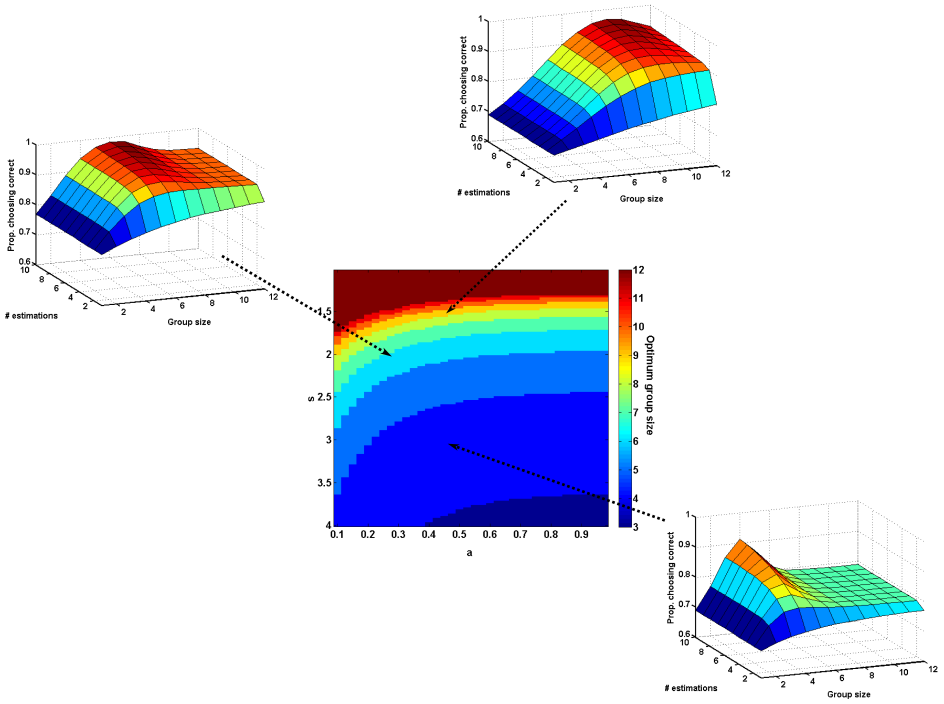

FIGURE 2.4: Prediction of the optimal group size depending the parameter combination. The presence and the value of the optimum depended entirely on the balance between the private and the social information. For very social animals, the amplification of mistakes at larger group sizes is going to be detrimental while the social influence can be positive if it is moderately low.

We also wanted to be sure that the model was general, and not a consequence of the formula of the decision rule 2.1. For this purpose, we tried to change the decision rule to other formulas that also used the social influence of other individuals for modifying the probabilities of choosing each option. We first used an extension of our decision rule, published in Arganda et al. (2012):

$$P_y = \left(1 + \frac{1 + a_y s^{-(N_y - kN_x)}}{1 + a_x s^{-(N_x - kN_y)}}\right)^{-1} \tag{2.2}$$

This formula has two main changes: (1) the parameter of the private information, a, has been substituted by two parameters, a_x and a_y, that are still related to the private information. This deals with the fact that individuals can now estimate how good each option is separately. (2) The new parameter, k, measures how

an individual choosing one option informs about the other option (see Arganda et al. (2012) for further information). We found that the presence of the optimum is conserved using this formula and, as happened with the previous model, the value of the optimal group size depends completely on the parameters we use (see Figure 2.5A). For some parameters, there is not even an optimal group size, due to additional effects of the decision rule.

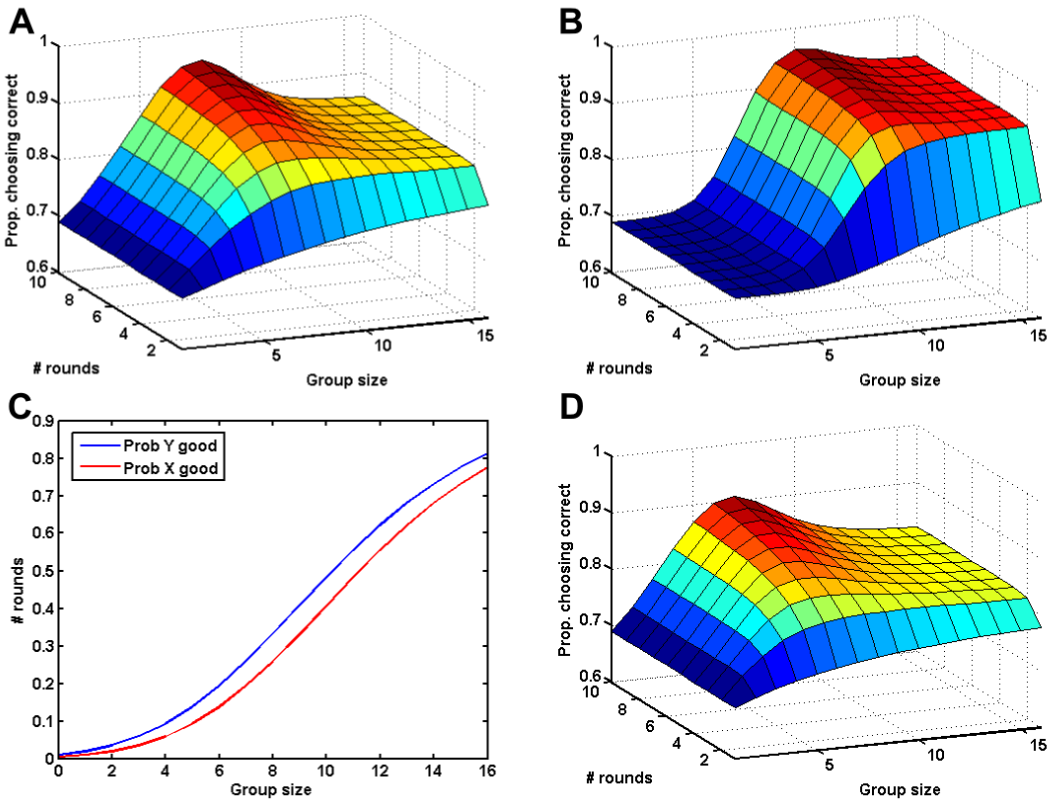

FIGURE 2.5: Prediction of an optimal group size for several decision rules. The parameters are chosen so that a single individual has the same probability of choosing the correct option as in figure 2.3. (A) Optimal group size for the formula in Arganda et al. (2012), with $a_y = 0.3497, a_x = 2, k = 0.7$ and $s = 2$. (B) Optimal group size for the formula in Sumpter and Pratt (2009), with $p_y = 1, p_x = 0.4499, a = 0.025, m = 1, k = 5$ and $T = 10$. (C) Optimal group size using a Hill Function as the deciding formula, with $y_{00} = 6.7814, x_{00} = 5.7738, K = 17$ and $n = 5$.

We then tried to use decision rules from other from other works and in this case we chose the formula from Sumpter and Pratt (2009). The probability of committing to option y for this model is

$$p_{y,good} = p_y \left(a + (m - a) \frac{N_y^k}{T^k + N_y^k} \right) \tag{2.3}$$

where p_y, a, m, T and k are parameters that define several properties of the function (see Sumpter and Pratt (2009)). A similar function determines the probability of committing to option x, and the probability of choosing each option is obtained by probability matching: $P_y = p_{y,good}/\left(p_{y,good} + p_{x,good}\right)$. We found similar predictions with this function, and again the presence of the optimal group size depends on the parameter combination (see Figure 2.5B).

The last formula we tried was a general mathematical formula, a Hill function, so the probability that option y is good is given by:

$$p_{y,good} = \frac{\left(N_y + y_{00}\right)^m}{K^m + \left(N_y + y_{00}\right)^m} \tag{2.4}$$

The shape of this function can be seen in Figure 2.5, and option x has its correspondent formula, with N_x and x_{00} instead N_y and y_{00}. Both x_{00} and y_{00} are related to the probability of choosing the correct option when no other individuals have decided and m and K are parameters of the model. The final probability of choosing option y will be computed as the previous model, by probability matching: $P_y = p_{y,good}/\left(p_{y,good} + p_{x,good}\right)$. The prediction of this model can be seen in Figure 2.5D, where we see that the presence of an optimal group size is still maintained. Therefore, the existence of an optimal group size for different decision formulas, provided that the social influence has a significant effect on the probabilities of the decision, seems to be a general consequence of our decision mechanism.

We then thought that the existence of an optimum may be caused by the precise decision mechanism of our model, so we decided to try other possible mechanisms to see if they still predicted the existence of an optimal group size. For this simulations we used again the original decision rule (formula 2.1). First, we tried to randomize the order in which individuals voted, so that the specific voting sequence was no longer conserved through consecutive rounds. This mechanism does not really change the previous predictions (Figure 2.6A), mainly because all the individuals have to vote at each round before reevaluating their opinion, so the information available at each point is not going to be so different from the original mechanism. In other words, in this mechanism the first voting individual has to wait for the votes of the rest of the group and, even though it may not be the first to vote in the second round, it will have access to much more information

than in the first round. This then creates a similar situation, compared to what we already had.

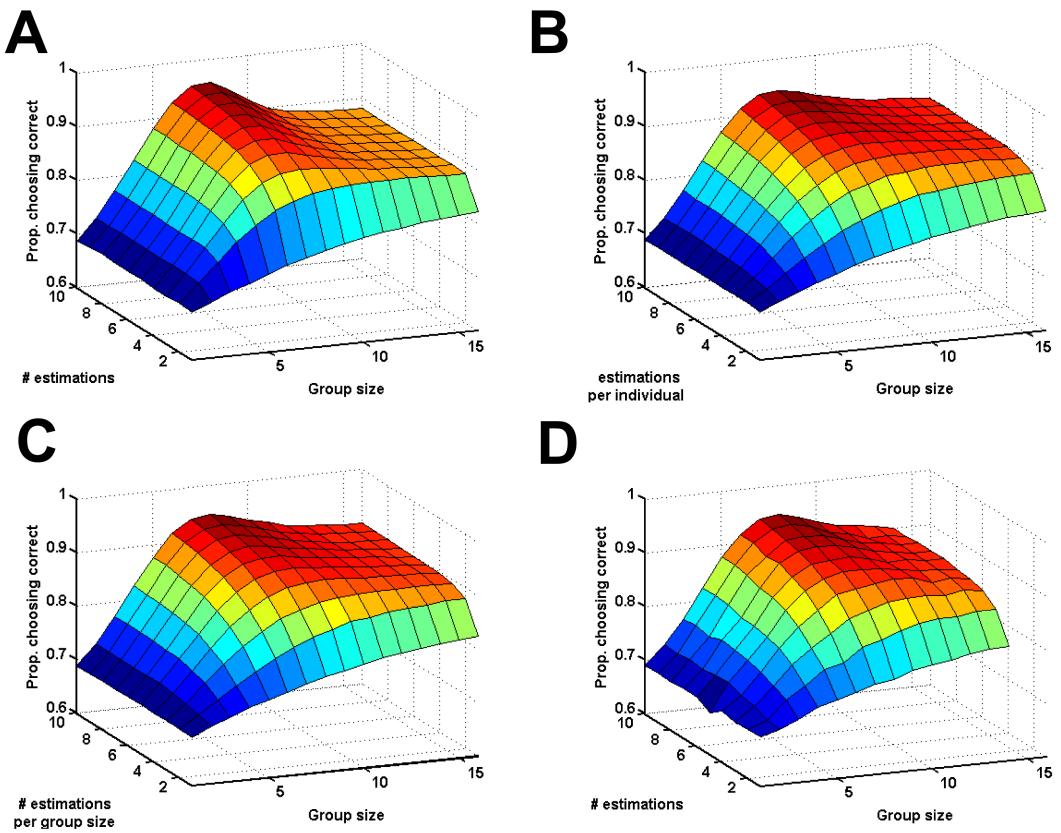

FIGURE 2.6: Optimal group size for different decision making mechanisms. The decision formula used is 2.1, with the same parameters as in Figure 2.3 (A) Randomized decision order without overlapping rounds. (B) Randomized order with overlapping votes between individuals. The order is no further conserved and each individual has a fixed number of votes. (C) Same mechanism as B, but now individuals do not have a fixed number of votes. (D) Mechanism with gaussian probability of voting for each individual. Each gaussian is centered on the individual and has a standard deviation of two individuals ($\sigma = 2$). At each position, the contributions of all the gaussians are added and normalized by the sum of the total area of the gaussians. This value gives us the probability of voting for each individual. Decisions are allowed to overlap and individuals do not have a limited number of votes.

For the next mechanisms, we are allowing votes to overlap, so the decision process becomes more dynamical. At each step, all individuals have a uniform probability of voting for an option, so we can no longer talk of sequences at the group level. For this case we have two subcases, one in which every individual has a fixed number of votes (Figure 2.6B) and another in which the number of votes is not fixed but we force every individual to vote at the end of the process if they have not done it before, so we can consider they have all decided (Figure 2.6C). These two models still predict the existence of an optimal group size, even though

the amount of randomization has increased. For the last model, each individual has a Gaussian probability of voting based on the order he occupies on the group (Figure 2.6D). This is an intermediate case between the ordered and the completely random model and it also predicts the existence of an optimal group size. The results of these section show that the existence of the optimum is a general fact of having several rounds of voting, independently of the mechanism or the rule of the decision.

2.2.4 Relation with Condorcet's Theorem

As our decision making model is completely general, we could calculate the probability that the majority of individuals choose the correct option and compare this outcome with Condorcet's Theorem. We find that the result for the first round is the general expected result, where Condorcet's Theorem predicts a better performance (see Figure 2.7A, left, blue and red lines respectively). However, this result does not hold for further rounds and, surprisingly, the accuracy of the decision surpasses that of Condorcet's Theorem for some group sizes (see Figure 2.7A, center and right). This is very significant, as the classical theory was only able to improve the accuracy of Condorcet's Theorem by using a subgroup of more informed individuals, whose votes counted more than those of poor informed individuals (Shapley and Grofman, 1984). In our model not only all individuals are identical, but also the weight of their votes is the same. Nevertheless, Condorcets Theorem still surpasses the performance of our model at large group sizes.

Condorcet's majority rule still represents an upper bound for the accuracy of collective decision-making in the case of a simple round (Sumpter and Pratt, 2009). This is the case because majority rule under independent voting maximizes the total amount of information of a whole group. The private information of each individual is shared and the voting cannot be negatively influenced by other members of the group. However, when individuals vote with some correlation among them, the accuracy of the decision is diminished because the total amount of information is the same as in the independent case but the mutual interaction can only contribute to low the quality of the decision, as copying behavior may amplify errors that do not occur in the independent case. In spite of that, when we let individuals interact in a second round, they are adding again private information to the collective decision, so the performance of the group increases. In fact, if we considered the case in which each individual at further rounds counted as a virtual

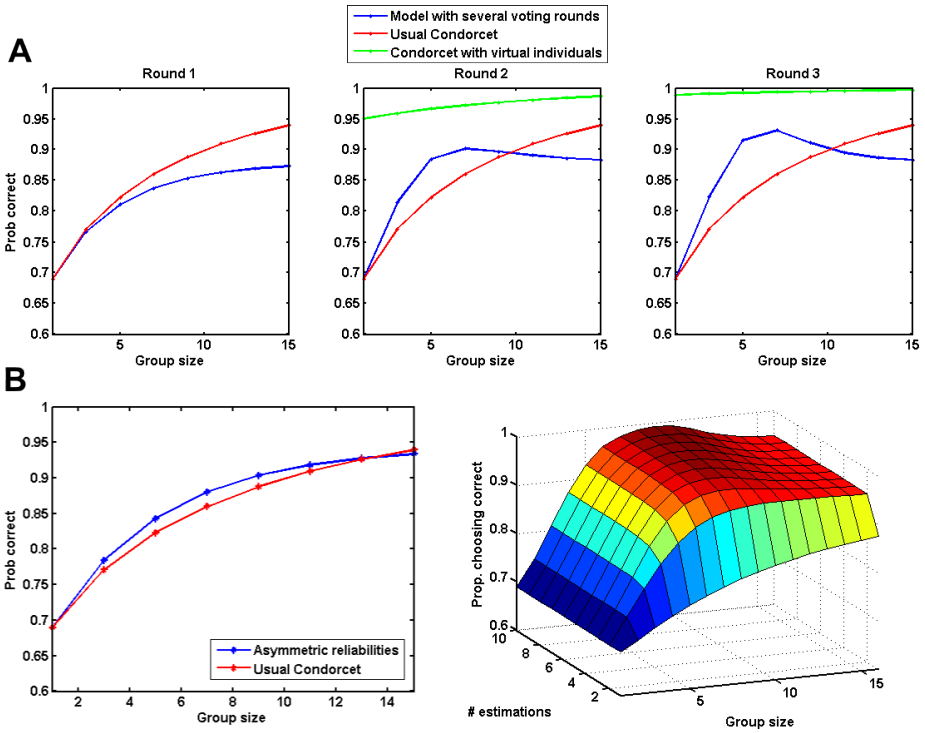

FIGURE 2.7: (A) Comparison between 3 rounds of the decision making model and Condorcet's Theorem. For simplicity we plot the results of majority voting for odd group sizes. At even group sizes the majority is not well defined at ties and they need to be settled by an additional rule. (B) Results from just one voting round using the decision rule with asymmetric reliabilities (left). The existence of an optimal group size is conserved using this decision rule (right). The parameters of the model used in formula 2.5 were $a = 0.45, s_x = 1.5$ and $s_y = 1.75$.

individual that votes independently, majority rule would again give a maximum in the performance of the group (see Figure 2.7, green line). However, this will never be the real result, as there is an anchoring effect in the individual answer when he is asked the same question repeatedly (Chen and Kemp, 2011; Mussweiler and Pfeif, 1991; Vul and Pashler, 2008). Interestingly, our decision making model presents a new mechanism that may solve the anchoring effect and that allows individuals to introduce more information in a collective decision.

Following this argument, we found that another way of introducing private information in the group would be to consider that the social information has more weight when the votes of other individuals agree with the private information of the deciding individual, so that the reliability of the individuals would be different depending on their vote (see Pérez-Escudero and de Polavieja (2011)). This implies an asymmetry in the reliabilities of other individuals that comes from the private information of each individual so, effectively, some extra private information is added to the decision process. It is an assumable scenario, as an individual may

trust more the individuals that agree with its own private information. Therefore, our decision rule will become

$$P_y = \left(1 + a \cdot s_x^{n_x} \cdot s_y^{-n_y}\right)^{-1} \tag{2.5}$$

where now s_x and s_y represent the reliability of the individuals going to option x and option y respectively (see Pérez-Escudero and de Polavieja (2011)). The results match our predictions and our model gives a better performance than Condorcets majority rule with just one voting round (Figure 2.7B, left). We have also checked that all the predictions from previous sections still hold for that decision rule, and the presence of an optimum is maintained (Figure 2.7B, right).

It is interesting to note that both ways of introducing extra private information do not improve the accuracy of Condorcets majority rule for large group sizes, and this means that the amount of information added to the decision does not overcome the negative effects of having social influence in these cases. It is evident that, for large group sizes, the total information handled by the group is so high that an independent voting system would always achieve the best accuracy in the decision.

2.2.5 Comparison to real animal data

So far we have seen that our decision mechanism has some relevant theoretical predictions, but we also wanted to confront it with experimental collective decision data to test its validity. As wee have seen in the introduction, there are several reported examples in nature where animals express their opinions about a specific choice in a collective decision (Sumpter et al., 2008, 2011; Ward et al., 2008b), so we decided to study the data of a collective decision experiment where one of the options was best than the other. We thus used the data from Sumpter et al. (2011), where different groups of mosquitofish (*Gambusia holbrooki*) had to decide between two options, in a Y-shaped set-up. The attraction for the two options was achieved making them deeper than the rest of the set-up. Additionally, a replica predator was placed at one of the options, creating a preference for the other option. We first tried to fit the data with our model of several voting rounds and the original decision rule 2.1, but it was not able achieve a performance as high as the data. However, if we added an extra amount of information by using two different reliabilities, the model could explain the data in an accurate way

(Figure 2.8A, red line). We also obtained a good fit for the data combining both the differences in the reliabilities and several rounds of voting (Figure 2.8A, blue line).

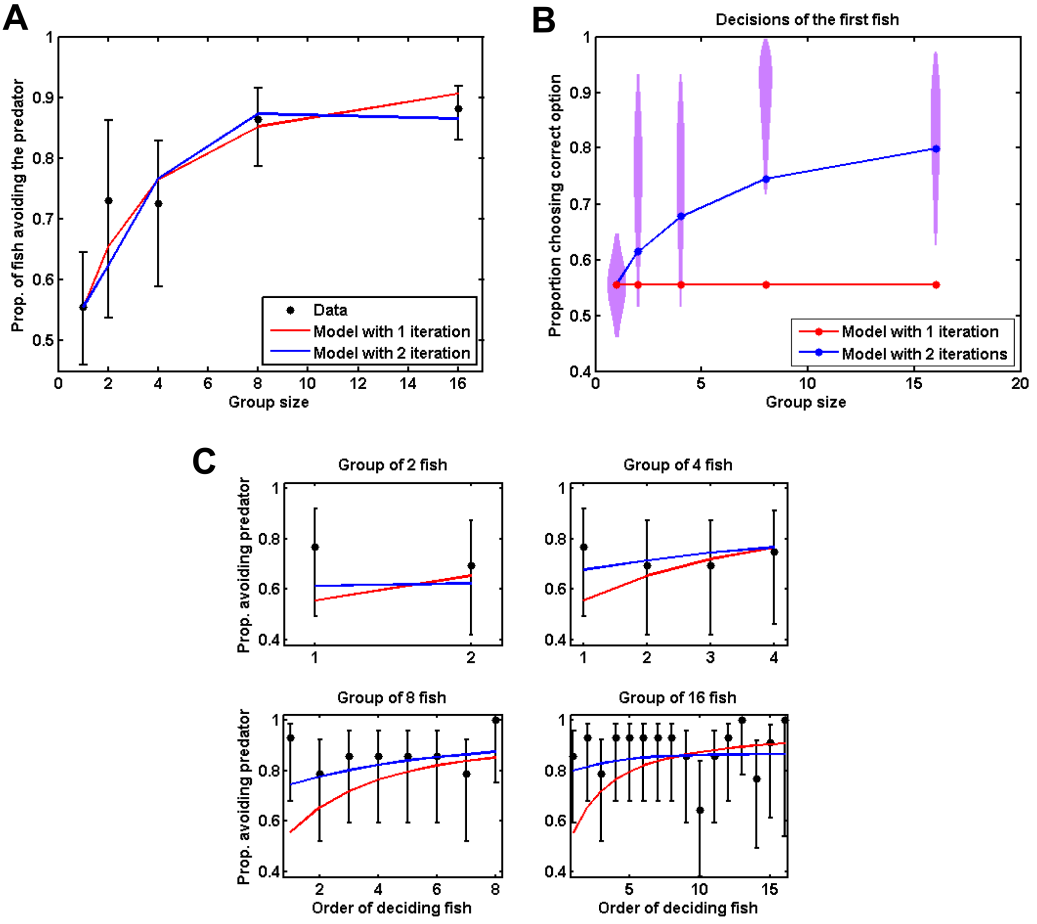

FIGURE 2.8: (A) Total proportion of fish choosing the correct option, for group sizes of 1, 2, 4, 8 and 16 individuals. Data is represented with black dots and the best fit is plotted for the models with one (red line) and two decision rounds (blue line). (B) Probability that the first fish chooses the correct option. Probability distributions showing the uncertainty of the experimental data are plotted in violet for each group size. The width is proportional to the probability of the real value and they are truncated at 95% confidence intervals. The results from the models of one and 2 rounds are also plotted, (red and blue lines, respectively). (C) Probability of choosing the correct option for each individual in the experiment. Outcomes from the models are again plotted with the same colors as before. The parameters of the model for this graphs are $a = 0.8$, $s_x = 1.55$ and $s_y = 2.4$. The best fit was obtained by calculating the root mean square error between the experimental data and the prediction of the model for figure A.

This result showed us that the two models can properly fit the data, so the necessity of several voting rounds still seems unclear. However, it became evident when we analyzed the data with a different quantitative approach. If we looked at the probability that the first deciding individuals in a group choose the correct

option, the two models gave very different outcomes (Figure 2.8B). We see that the proportion of first deciding fish choosing the correct option increases with the group size. This is something that sequential models with just one vote per individual cannot explain (see red curve) because, as there are no individuals that have previously chosen any particular option, they predict that this probability is constant and independent of the group size. If we look at the model with two voting sequences, we see that it is able to explain these data very accurately (blue curve). Furthermore, if we analyze the complete decision sequence, the model with two rounds explains the data much better than the model with only one (Figure 2.8C). These last two results argue in favor of the models where animals can reevaluate their initial preferences, as they clearly show that animals share information in a collective decision making process. They also point out that a partial analysis of this kind of experiment can be simplistic or misleading, as we have seen that a sequential model with no additional estimations can explain certain data but not if we look at them deeply.

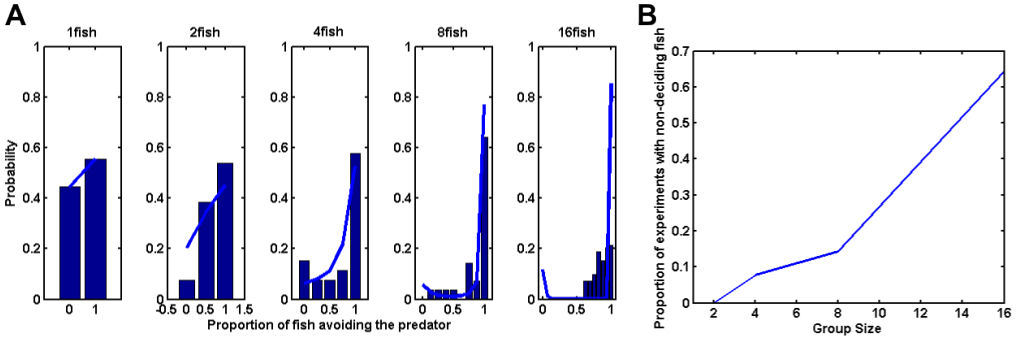

FIGURE 2.9: (A) Probability of the different outcomes for all group sizes. Blue line is the prediction of the model after two rounds. We used the parameters of the best fit to the data (Figure 2.8). (B) Proportion of experiments with at least one non deciding individual. We see a clear increase for the group of 16 fish.

We obtained the best fit for the model with just two sequences of voting. If we do the calculus, we find that the optimum for this parameter combination is obtained for a group size of 8 or 9 individuals depending on the number of sequences. This suggests that a bigger group would have a worst performance in this experiment, and there are some facts in the data that point into that direction. If we look at the behavior of the whole group we find that for a group size of 4 or 8 fish, the majority of the experiments end with a consensus where all the fish go to the correct choice (Figure 2.9A). Surprisingly that does not happen for the group of 16, where the group of fish usually splits and we only see consensus in 20% of the experiments. Furthermore, in a high proportion of cases we found that there were

fish that did not make a final decision or its behavior was not clear (Figure 2.9B), suggesting that the group size was too big for this specific experiment.

2.2.6 Extending the model to Self Propelled Particle Simulations

One of the weakest points of our decision making process is its lack of dynamics because (1) a group of deciding fish will probably not update their votes or opinions in such a deterministic way, and (2) the experimental setup imposes certain restrictions about the number of times the voting process can occur. To overcome these difficulties we decided to use a more dynamical approach, integrating our model in individual-based simulations of particles moving in the plane. Works like Couzin et al. (2002, 2005b) or Nabet et al. (2009) used a similar approach, studying how discrete simulations of particles based on individual rules could created a common pattern or even a collective decision.

For the simplicity of our simulations, the setup was a rectangular arena with two alternative options at the end of one of its major sides (see Figure 2.10A). At the beginning of the simulation, animals moved towards the two possible options by choosing randomly one of them. Then, each animal had a fixed probability per iteration of voting for one of the options while they advanced towards the end of the setup. As we are trying to explain the previous experiments of decision making, we decided to use the same decision rule that fitted the data (formula 2.5). Therefore, every time that a fish made a decision by chance, he evaluated the two options using this formula and chose one of them. The simulation finished when all the fish have reached one of the decision zones. Interestingly, the simulations gave a good fit for the total proportion of individuals (Figure 2.10B), but a surprisingly accurate fit for the first deciding fish and for the sequence of decisions (Figure 2.10B and 2.10C, respectively). The fit was even better than the one we obtained in Figure 2.10, something that argues in favor of the use of more dynamical models.

In this sense, we decided to create a new model of particles where animals did not have to choose where to go between two options, but instead they could choose any point of the setup, something that seems closer to real experiments. We thus binarized the space in pixels and gave each of them a probability of being chosen by each individual at each iteration (see Figure 2.11A). The probability of going to pixel i was given by:

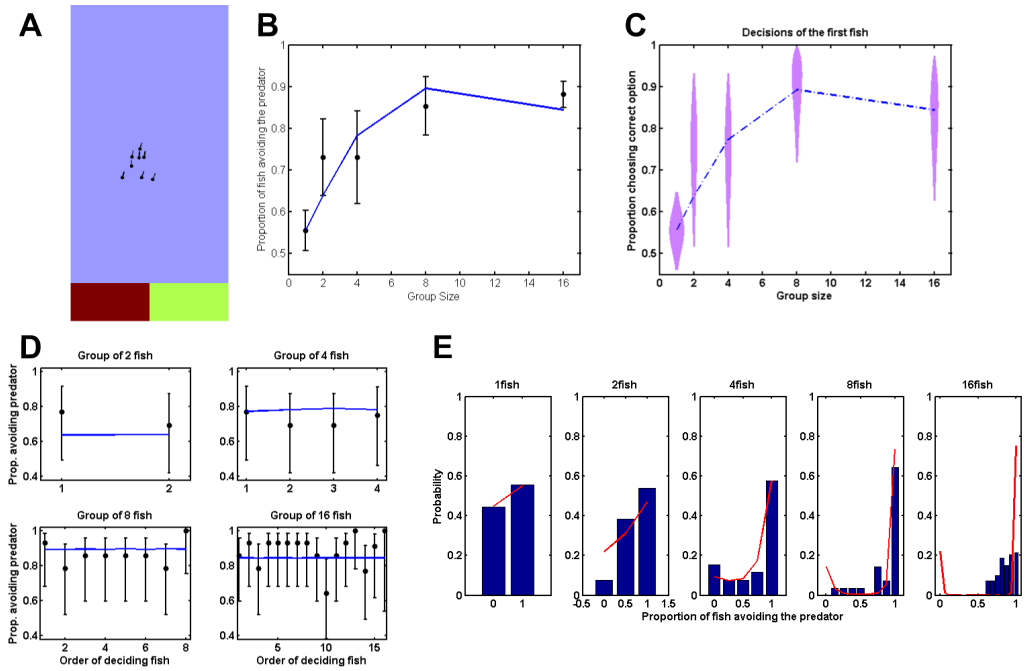

FIGURE 2.10: (A) An example frame of the simulation. (B) Total proportion of fish choosing the correct option. Data is represented by black dots and the model is the blue curve. The parameters used in the decision rule (formula 2.5) were $a = 0.8$, $s_x = 1.45$ and $sy = 1.75$. (C) Probability that the first fish chooses the correct option. Data are represented in violet as in Figure 2.8B while the model is represented by the blue dashed line. (D) Sequence of the decisions as in 2.8C. (E) Probability of the different outcomes, as in Figure 2.8D.

$$P_i = \left(1 + a_i \cdot s^{-N_i}\right)^{-1} \tag{2.6}$$

where a_i measured the quality of nonsocial information of pixel i, s measured how reliably an individual at pixel i indicates that this pixel is a good option and N_i is the number of individuals on that pixel (see Arganda et al. (2012)). The nonsocial quality of one of the deciding options was better than the other, creating the desired asymmetry. Still, both of them were better than the rest of the setup, where we created a gradient that improved the quality of the pixels depending on the proximity to the deciding zones. This gradient motivated individuals to move towards the deciding zones, as it happens with the depth of the setup in the real experiment. We gave individuals a certain radius of influence, as punctual individuals affecting only one pixel would not create the desired effect. Apart from that, a given individual could indicate that the location towards it is heading (rather than its current location) is a good place. To account for this, we centered the circle of influence of individuals at a position where it would be in the future , assuming it would keep constant direction and speed (t_{future} gives the number

of steps to calculate the future position of the individual). Finally, N_i is the number of individuals that have social influence on pixel i. At every iteration each individual has a fixed probability $p_{decision}$ of choosing a new pixel and it will

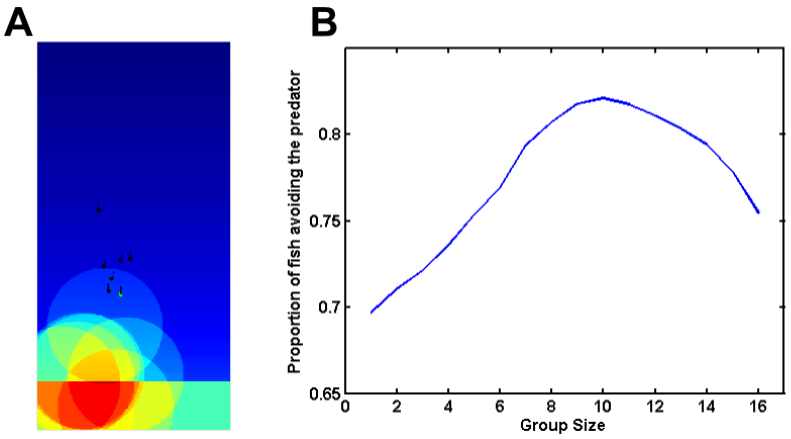

FIGURE 2.11: (A) An example frame of the simulation. (B) Proportion of individuals choosing the correct option as a function of group size. The parameters used were $p_{decision} = 0.2$, $s = 1.65$ and $t_{future} = 15$. For this model, the higher the values of a, the worst this pixel become, so the value of a the correct option was $a = 105$ and for the bad wrong was $a = 194$. The gradient was created linearly from top ($a = 1000$) to bottom, or the end of the setup ($a = 10000$). The curve is the average of 10000 simulations

The outcome of this model confirmed that this specific spatio-temporal simulations produced also an optimal group size. In figure 2.11B we clearly see that the proportion of individuals choosing the correct option has a maximum at a group size of 10 individuals. This was the last step supporting that the presence of an optimum is inherent to a social decision making process that involves the interaction or votes between individuals and the possibility of change of opinion.

2.3 Discussion

In this chapter we present a new decision making model in which individuals are allowed to reevaluate and change their opinions. This model predicts the existence of an optimal group size by balancing the benefits of receiving social information with the negative effects of amplifying initial mistakes. This prediction has been shown to be a general aspect of this kind of models, as it is maintained in spite of the addition of several modifications to the original model. We have also shown that our model improves the accuracy of Condorcet's Majority Rule, a decision making process that was assumed to be an upper bound of collective decisions. Moreover, our model has successfully been confronted against experimental data, supporting its validity.

A very different model that also fitted real data and dealt with optimality was proposed in Amé et al. (2006). In their research, they studied a two option case of site selection in cockroaches. They saw that the probability of leaving or joining a shelter depended on two parameters: its capacity and the number of individuals already there. With these data, they built a model and predicted the existence of an optimum in the benefits of being under a shelter, depending on the combination of the previous parameters. This study is especially appealing because it operates in a similar way as our work, obtaining theoretical predictions of a model that fits experimental data in a specific case. The contribution of our model is a further improvement, as it may be applied to collective decisions in general, not to a specific case of shelter selection.

Our model opens a new way to explain collective decision making processes by including individual changes of opinion. It is particularly interesting that it can explain details that are not commonly studied by other models, such as the probability of choosing the correct option depending on the order in which the individual decides. In general, decision making models are only designed to partially explain the experimental results, focusing only on a specific outcome of the data, such as the total proportion of individuals choosing the correct option (Sumpter et al., 2011; Sumpter and Pratt, 2009; Ward et al., 2008b). This is the case for the latter paper, where the authors used a theoretical model to exclusively explain the results of the total proportion of fish avoiding the predator. In their model, the group will avoid the predator if at least one of the individuals detects its presence, whereas they will choose randomly if the predator is not detected. In spite of its

simplicity, the model fits quite accurately the data they used. However, if we face the model against all the possible final configurations (see Figure 2.12A), we find that it predicts a peak at a 50% proportion that is not present in the experimental results. Thus, in order to better understand the implicit process of collective decisions, we claim that a bigger effort should be done in analyzing more extensively different parameters of the experiments, including dynamical ones.

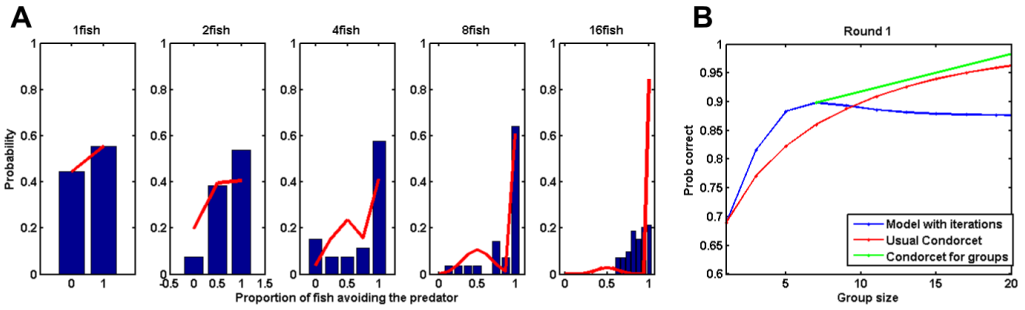

FIGURE 2.12: (A) Prediction of the many eyes model (red line) for the different outcomes of the experiment (blue bars). (B) Effect of combining our model and Condorcet's rule for the outcome of groups of discussion. For a total number of 21 individuals: prediction of Condorcet's Theorem (red line), prediction of our model for two voting rounds (blue line) and Condorcet's rule applied to three different groups of 7 individuals that have previously voted following our model.

Our results also points out that the main cause of the existence of an optimal group size is the amplification of the decisions made by the first individuals. Since there is a low level of social information at the beginning of the process, these first individuals will have less amount of total information, so their accuracy will be limited. If, by chance, the first individuals tend to choose the wrong option, their initial mistakes may be amplified, specially in large groups. This amplification would not happen in smaller group sizes, as the first deciding individuals have the chance to reevaluate their opinion soon enough in the process. Groups of intermediate sizes find a compromise between both things, as they do not amplify so dramatically the initial mistakes, but at the same time they incorporate enough amount of social information. This effect was described in Kao et al. (2014), where they use a very ambiguous model assuming that animal decisions depend on environmental cues that point at the correct option with different reliabilities. Even though the idea seems quite natural, the implementation of the model is not totally clear and seems to be quite unrealistic. Our model, however, creates an ideal confluence between their arguments about the accuracy obtained by intermediate groups and a model that could be further used to explain different experimental collective decisions.

The existence of an optimal intermediate group size in decision making is specially appealing as it competes with the idea of the Wisdom of the crowd. This theory claims that collective estimations improve upon the estimations of most individuals of a group (Galton, 1907; Lee and Shi, 2010; Surowiecki, 2004; Wagner and Vinaimont, 2010). In these theories, as in Condorcet's Theorem, individuals always make independent estimations and they are all pooled together afterwards. The need for independence argued against ideas involving interactions between the deciding subjects, such as democracy or groups of discussion. So far, the improvements in the accuracy of a collective decision was achieved by including a minority of more informed individual whose votes weighed more (Shapley and Grofman, 1984), resembling the situation of oligarch governments. Furthermore, it has always been considered that the interaction between individuals lowers the accuracy of independent estimations pooled together (Berg, 1993; Ladha, 1992; Nitzan and Paroush, 1984). Thus, our model, seems to be the first theoretical evidence defending that discussion groups may be useful for solving collective problems. As we mentioned before, it is able to do so because individuals aggregate more information with each voting round while they are able to pool it due to the option of change of opinion. Moreover, this will also prevent the anchoring effect observed when individuals are asked the same question repeatedly (Chen and Kemp, 2011; Mussweiler and Pfeif, 1991; Vul and Pashler, 2008). In the theoretical limit of our theory, the best accuracy would be achieved by dividing a group in several subgroups and applying Condorcet's Majority rule at the outcome of the subgroup discussions (see Figure 2.12B).

Another elegant and novel aspect of our model is that it predicts the existence of an optimal group size without having being designed for it. Generally, studies trivially predicted the existence of an optimal or stable group size by having a balance between advantages and inconveniences of living in group (Brown, 1982; Higashi and Yamamura, 1993; Sibly, 1983). In this sense, the prediction of our model is complementary to the model itself, highlighting its importance. We are not claiming that animal group sizes observed in nature can be explained with our decision making model, as this process is far more complicated. Furthermore, the existence of an optimal group size is not observed globally and some studies even claim that animals group sizes follow a power law distribution (Bonabeau et al., 1999; Niwa, 1998, 2003). However, their work has only been applied to pleagic fishes and a few groups of herbivores, so the differences in the environmental conditions of other species may produce different results. It could also be the case

that animals live in a specific group size during a definite amount of time in order to optimize their performance in a specific task. These tasks, such as hunting, breeding, or even making collective decisions, may have its influence on the group size. We then claim that our decision making model could have important theoretical predictions that could be appealing for the scientific community. A strong point is that these predictions can be tested experimentally, both in animals and in humans, creating further insight on the mechanisms of social decisions. It should also be considered for explaining more general experiments of collective decision making, as it presents a new way of integrating dynamical information transfer between individuals, a key aspect in group decisions.

2.4 Conclusions

- We developed a new model of decision making in several voting rounds that improves the accuracy of collective decisions by increasing the amount of information of the first deciding individuals. This individuals generally decide with less social information, so several voting rounds and the option of opinion change improve their accuracy.

- Our model predicts the existence of an optimal group size in decision making tasks. This prediction is caused by the fact that intermediate group sizes do not experience the difficulty of large groups at correcting individual mistakes while they still benefit by receiving enough social information. Furthermore, this prediction holds for several models and several decision rules, as long as they allow individuals to reevaluate and to change their initial opinions.

- We showed that our model can improve the accuracy of Condorcet's Theorem, something that was not previously achieved with non-independent voting models. The introduction of more private information by the votes of the individuals at each round seems to be essential for this result.

- We proved that our model can explain the results of an experimental data set, confirming its validity. Our model was able to fit different parameters of the same experiment, something that has not been fully studied by previous collective decision making models.

- We showed that two models of self propelled particles deciding between two options also predict the existence of an optimal group size. This result further confirmed the prediction of an optimal group size is inherent to models that include change of opinion.

2.5 Conclusiones

- Hemos desarrollado un nuevo modelo de toma de decisiones que permite a los individuos votar varias veces y cambiar de opinión al enfrentarse a una elección entre dos opciones. Este modelo mejora la decisión a nivel de grupo por el hecho de aumentar la cantidad de información que le llega al primer invididuo, en general más desinformado.

- Nuestro modelo predice la existencia de un tamaño óptimo de grupo para la toma de decisiones. Esta predicción se debe a que los tamaños de grupo intermedios no presentan tantas dificultades como los grupos de mayor tamaño a la hora de corregir errores iniciales mientras que se siguen beneficiando de la información social que reciben. Esta predicción se mantiene aunque cambiembos la fórmula que los individuos usan para tomar su decisión, siempre y cuando se les permita reevaluar y cambiar sus opiniones iniciales.

- Demostramos que nuestro modelo puede mejorar la precisión del Teorema de Condorcet, algo que no se había conseguido hasta ahora utilizando modelos en los que los votos de los individuos tuvieran algún tipo de correlación. La manera de introducir información privada mediante los distintos votos en cada iteración parece ser determinante para este resultado.

- Probamos que nuestro modelo es capaz de explicar un conjunto de datos experimentales, confirmando su validez. Nuestro modelo explica distintos parámetros del mismo experimento, algo que no se ha intentado hacer de forma general con anteriores modelos de toma de decisiones en grupo.

- Mostramos que dos modelos diferentes basados en simulaciones de partículas también predicen la existencia de un tamaño óptimo de grupo. Este resultado sirve para confirmar que la predicción de un tamaño óptimo de grupo es inherente a modelos de toma de decisiones que permiten el cambio de opinión en los individuos.

2.6 Materials y Methods

2.6.1 Computing the final probabilities

Instead of performing numerical simulations, we did the exact calculus for the final probabilitie in all the cases where it was possible. The first voting round was simple, as we just had to build the tree of probabilities, using the corresponding decision rule at each point (see Figure 2.13 for the case of two individuals). The possible decisions of the individuals are going to build a tree of states that reflect what individuals have chosen. The probability of each state will be directly the product of the probabilities that lead to it. However, for further iterations, all the states already exist so, after an individual's decision, each state would lead to two existing states. This implies that the final probabilities of each state will be the sum of the probabilities of all possible pathways that lead to it. The states represent, at each point, the decision of all the individuals. Thus, in order to calculate the histograms or the total group probabilities we just had to multiply the probability of each state by the proportion of individuals it represents, and then just add the probabilities of the states that represent the same number of individuals at each option.

This algorithm of calculating probabilities could only be applied when all the individuals made a decision in each voting round, so it could not be applied to the cases where we let the decisions between the individuals overlap. In these cases, we simulated the decision process and then averaged the results over 10000 to 100000 simulations, depending on how noisy they were.

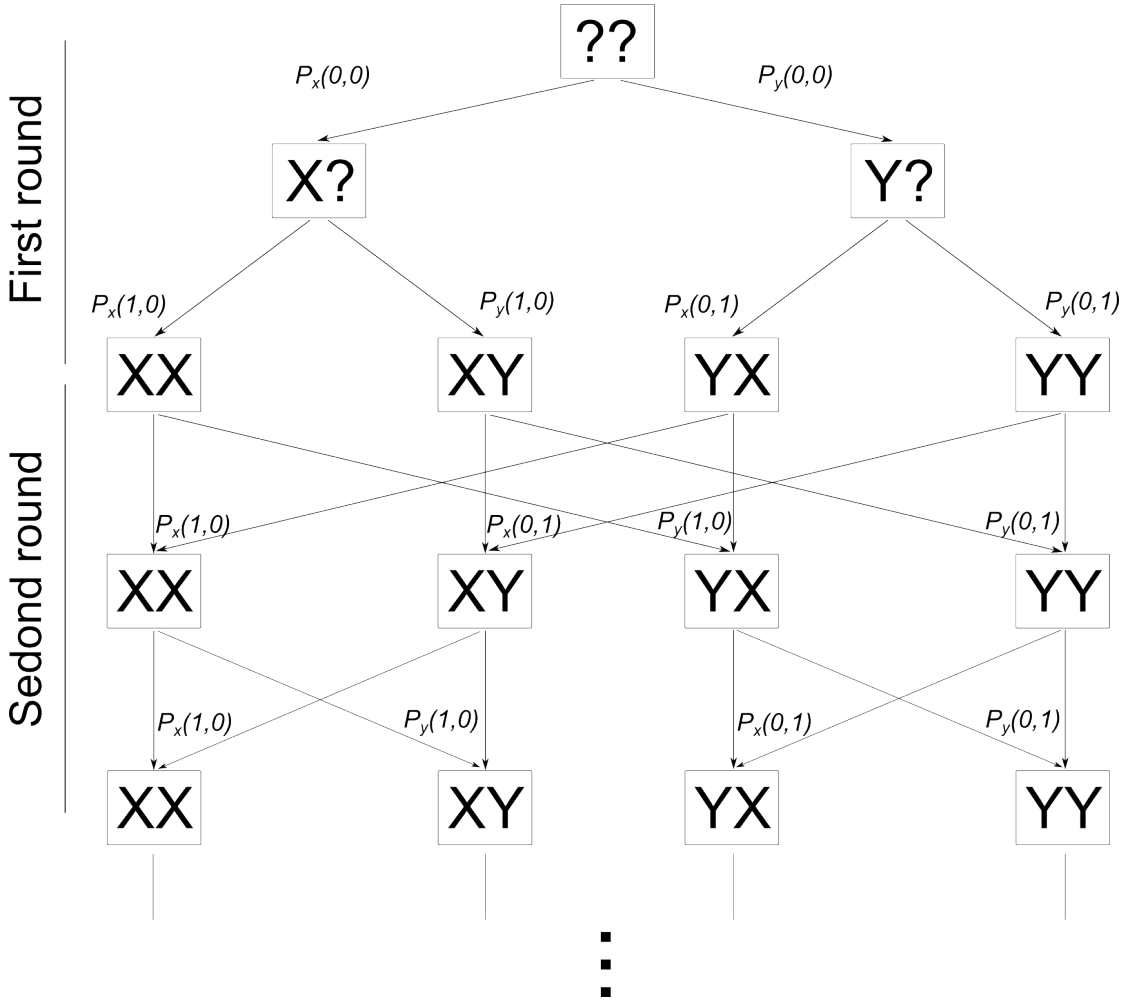

FIGURE 2.13: Building of the probability tree for a case of two individuals. The probability of each decision is represented by $P_{x,y}(N_x, N_y)$, where x or y represent each of the options and N_x and N_y are the number of individuals at option x or y, respectively. In the first round, the first individual to decide does not have any social information ($N_x = 0$ and $N_y = 0$) and there are only two possible states after this decision, one with the first individual at site x and the other with the individual at site y. For the second individual, now N_x or N_y will be different to zero, and this has to be included in the calculus. For the second round, both individuals have already decided, so each state leads to two already existing states, a path showed by the lines that connect the states. As individuals don't count themselves, when the first individual decides again it only sees what the second one has done and thus $N_x = 1$ or $N_y = 1$. The second round will continue with a similar situation for the first individual. At the end of it, there are several entangled paths that lead to each state and its probability would be the sum of all these paths.

3. Epigenetic modulation of behavioral individuality

3.1 Introduction

The topic of this chapter is going to be phenotipic variability between the members of a population. Classically, the phenotypic diversity has been considered to be generated by the genetic differences between the individuals and the disparity of their environmental influences (Galton, 1874). In other words, it has generally been accepted that phenotypic diversity within a population is produced by differences in alleles and differences in environmental inputs that modify gene expression (Jaenisch and Bird, 2003). According to this idea, isogenic individuals kept in the same narrowly controlled environment should be phenotypically identical, which has been shown not to be the case, as there are some studies that show the existence of a residual variability aside of genetic and environmental variation (Gärtner, 1990; Vogt et al., 2008).

The existence of this third source of phenotypic variation has been proved to exist in an extremely wide range of species, from viruses and bacteria to plants or animals (Vogt, 2015). Particularly in animals, a clear example are the morphological differences that exist in monozygotic twins (Jain et al., 2002; Seidel et al., 2003) and even in clonal animals (Shin et al., 2002). As a further example, more in line with this thesis, recent works have also shown the existence of behavioral variability independent of genetic differences in flies (Kain et al., 2012) or in mice (Freund et al., 2013).

There are several causes that might contribute to this type of variability. One of them is the existence of stochasticity in many biochemical processes in the cell that could lead to differences in gene expression (Elowitz et al., 2002; Kaern et al.,

2005), random cell fate determination (Nanjundiah and Bhogle, 1995) or variable morphogenetic signalling (Nijhout et al., 2003; Veitia, 2005). Other causes are more related to different experiences that arise from interacting with the environment (Freund et al., 2013). For instance, individual decisions about the amount of food to eat could create slight differences that could get amplified with growth, probably resulting in different ways of allocating metabolic resources. We cannot discard other mechanisms for the generation of variability, such as maternal or paternal effects, that could be transmitted to the offspring (Seong et al., 2011).

Even though the knowledge about behavioral variability independent of genetic differences has increased substantially, the molecular machinery that participates in this behavioral individuality is still unclear. Neuronal mechanisms such as neurogenesis, or serotonin signaling have been shown to be final targets of individuality in different models (Freund et al., 2013; Kain et al., 2012), but the pathways required to enable behavioral variability remain elusive.

One of the main process that could mediate the production of different phenotypes from the same genotype is chromatin modification. There are two main mechanisms that modify chromatin: DNA methylation and histone modifications, and both can be collectively referred as epigenetic mechanisms. This chromatin modifications can mediate the interactions between histones and DNA, and help determine whether DNA is accessible for gene transcription (Strahl and Allis, 2000). They can either be stochastical or a consequence of different environmental signals, but they will eventually result in the generation of phenotypic variability (Jaenisch and Bird, 2003; Weigel and Colot, 2012).

The capacity of encoding stable differences between individuals make chromatin modifications a strong candidate for being the molecular mediator of variability. In this sense, studies using monozygotic twins have demonstrated that epigenetic variations are accumulated during development (Fraga et al., 2005) and that chromatin changes can mediate the effects of the paternal diet on the offspring (Öst et al., 2014). In this chapter we will focus histone acetylation, whose role in behavior has been shown using different experimental models (Heller et al., 2014; Wang et al., 2013) with the additional participation of transcription factors driving the process (Whitney et al., 2014).

As a conclusion, we reasoned that molecular substrates linked to chromatin modifications could lead to differences in individuals by the generation of specific stable epigenetic and transcriptional profiles relevant to behavior. Our aim then was to

study the molecular mechanisms required for the generation of behavioral individuality independent of genetic differences under uniform environmental conditions. For that purpose, we chose zebrafish for its advantages in large-scale behavioral analysis, its wide genomic information and the simplicity of its pharmacological treatments.

3.2 Results

3.2.1 Behavioral individuality in larval zebrafish is stable for days

We decided to study behavioral variability in larvae from 5 to 8 days post-fertilization (dpf) as they already display a rich repertoire of sensory-motor behaviors (Brockerhoff et al., 1995; Kimmel et al., 1974; McElligott and O'Malley, 2005), accompanied by neurogenesis and the establishment of stable neural circuits (Burrill and Easter, 1994; Roeser and Baier, 2003; Thirumalai and Cline, 2008). As we wanted to study behavioral individuality, we built a setup in which we could record as many larvae as possible. The setup consists of a camera pointing downwards over two multi-well plates, where larvae were placed (see figure 3.1A). There were 24 circular wells for each plate, so we could record 48 larvae at the same time. The wells are carved on transparent PMM and have their walls tilted so that even in the most lateral wells the wall never hides the larva from the camera. We used infrared light to have a controlled homogeneous amount of light that did not heat too much the water (see section 3.6.3 for additional details).

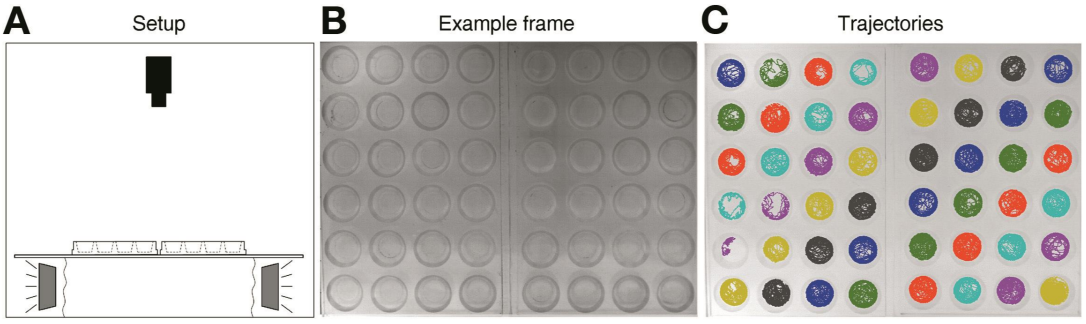

FIGURE 3.1: (A) Experimental setup. An overhead camera records the wells from above, with infrared illumination. (B) A single frame from an example video. (C) Reconstruction of the trajectories of the same video.

We first tested that zebrafish larvae at this stage are suitable for the study of behavioral individuality by recording them swimming freely in a single circular well for twenty minutes. We used the software multiwellTracker (developed in the lab and downloadable from www.multiwelltracker.es) to extract the trajectories of 48 simultaneously recorded individuals (see figure 3.1B for a frame of an example video and figure 3.1C for the reconstruction of its trajectories).

We chose overall activity (percentage of time in movement) and radial index (average relative distance from the border towards the center of the well) as parameters for representing individuality in larval zebrafish. We first proved that these individual parameters were stable over days, a necessary condition for talking of individuality. In figure 3.2 we plot the trajectories of the same population recorded from 5 to 8 dpf and we can see that they seem to be stable over days. To confirm that, the average of each individual parameter was tested from 5 to 8 dpf using Pearson coefficient of correlation (see Figure 3.3). Permutations of parameters between individuals were generated to obtain a P-value, giving R2=0.48, P<0.001, and R2=0.41, P=0.011 for activity and radial index of 5 dpf vs. 6 dpf respectively; R2=0.62, P<0.001, and R2=0.49, P<0.001 for 7 dpf vs. 8 dpf; and R2=0.69, P<0.001, and R2=0.58, P<0.001 for 7 dpf vs. 8 dpf.

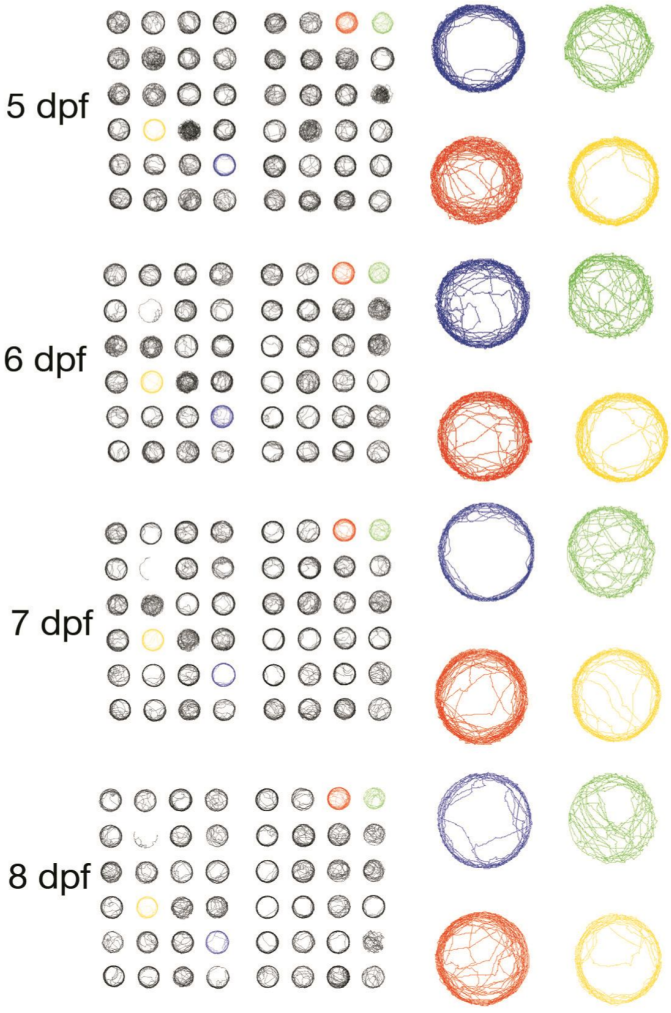

FIGURE 3.2: Right: Trajectories of the same population from 5 to 8 dpf. Left: Enlarged trajectories of 4 fish from these population. Each color represents the same larva during the complete experiment.

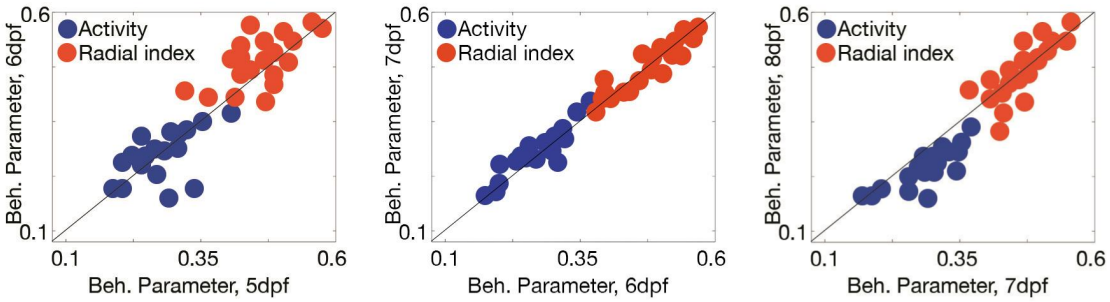

FIGURE 3.3: Correlation of activity and radial index throughout days, from 5 to 8 dpf (from left to right).

We first discarded the dependence of this spontaneous behavior on the position of the well in the setup. There were no differences in the individual parameters when we rotated the plate (90 degrees counterclockwise), or even when we interchanged the positions of larvae from outer to inner wells and viceversa (Figure 3.4A left, R2=0.73, P<0.001, and R2=0.68, P<0.001 for activity and radial index of a rotating plate; and 3.4A right, R2=0.65, P<0.001, and R2=0.61, P<0.001 for activity and radial index of interchanging wells). As another control, we found no correlation between the small differences in illumination across wells and the behavior of the larvae (see Figure 3.4B). Furthermore, using data from different recording groups of zebrafish, we did not obtain any significant difference in the parameters of unrelated individuals placed in the same wells (Figure 3.4C).

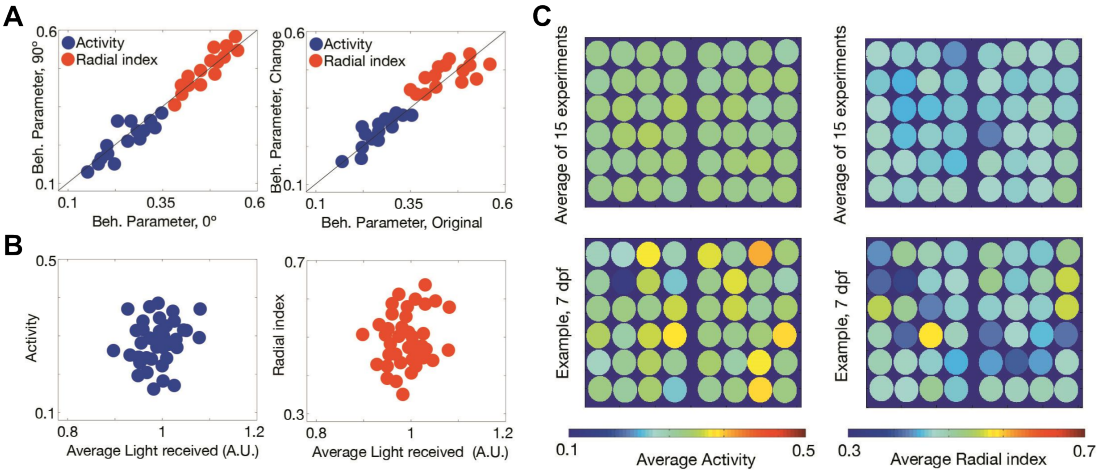

FIGURE 3.4: (A) Correlation of activity (blue) and radial index (red) after 90 rotation of the plates (right) and after interchanging larvae from external to internal wells (right). (B) Larval behavioral parameters (activity, left; radial index, right) plotted against the average light received by these larvae. (C) Average activity (left) and radial index (right) of a pool of different larvae recorded in the same wells from 15 experiments, compared to the same graphs using a single population (the test group at 6 dpf)

We still needed to establish that our setup could be used to study variability across individuals. In order to combine the two behavioral parameters we measure, we built a two-dimensional phenotypic space whose axes were our behavioral parameters and where we could represent the behavior of every individual and its intra-individual variability. For that purpose, we divided each video in consecutive fragments of 30 seconds and, for each fragment, we calculated the activity and the radial index of each individual. We then fitted the values of each individual to a two-dimensional Gaussian distribution and chose an isocontour of it for representing the behavior of each individual (Figure 3.5 left column). An isocontour is an ellipse with principal axes given by the eigenvectors of the covariance matrix of its values, and we chose them because each individual behavior can simply be represented by an ellipse on our phenotypic space. We chose the isocontour with length of each semiaxis given by the square root of the eigenvalue of the covariance matrix, as this reduces to the standard deviation in each direction for cases with no correlation between the two variables.

It is apparent from these graphs that the intra-individual variability, depicted as the size of the individual ellipses, is smaller than the inter-individual variability. In order to represent this latter value, we used the probability density of finding an individual with a given mean activity and radial index (Figure 3.5, right column. See section 3.6.3 for details of the smoothing procedure). This probability density provides a way to visualize behavioral variability in the population by the height of its peak or by its variance. We used *generalized variance* **(Wilks, 1932)** as a single parameter summarizing the two-dimensional variability in the population. *Generalized variance* is calculated directly from the raw data as the determinant of the covariance matrix, and has an appealing interpretation as measuring the surface of the ellipse fitting the data. Our statistical tests used *generalized variance*, but other variability parameters like the standard deviation of activity and radial index separately gave similar results **(Table S1)**.

We wanted to confirm in a quantitative way that the intra-behavioral variability was smaller than the inter-behavioral variability of the population. For that purpose, we built distributions that represented the intra-variability and the inter-variability of the activity and the boldness of a population separately. Intra-individual variation distributions were calculated using the coefficients of variation (CVs) of the 30 second fragments for each individual. In order to represent the complete variability of the population, we calculated the CV of the behavioral

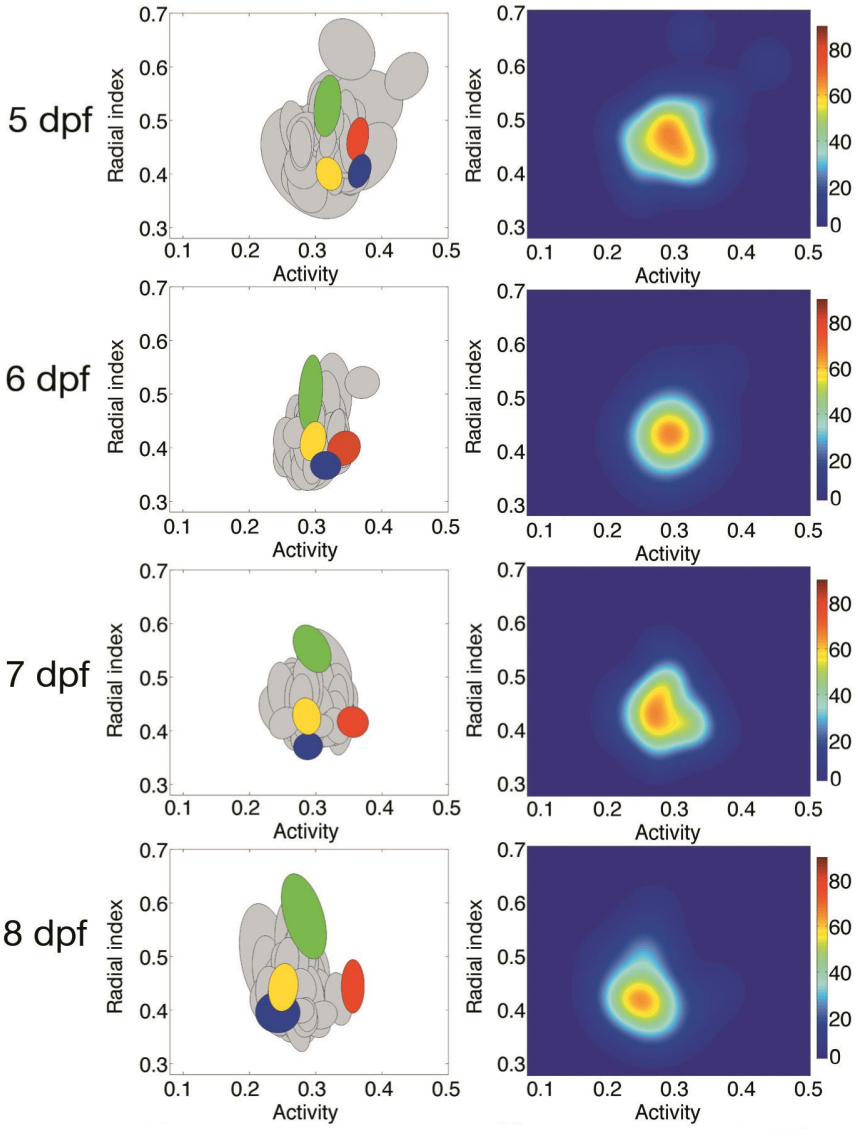

FIGURE 3.5: Left column. Population variability in activity and radial index of the same group from 5 to 8 dpf. Each ellipse represents the behavioral variability for each fish as the isocontour of the Gaussian fitted to the data. Colors mark the animals highlighted in figure 3.2. Right column. Probability density of finding an individual with a given mean activity and radial index.

parameters of all the individuals for each time fragment. Inter-individual distributions were built with this CVs, and the comparison between the smoothed histograms of the intra-variability and inter-variability distributions are represented in figure 3.6A . We see that the intra-individual variability is consistently smaller than the inter-individual variability ($P<0.001$ in all cases). Moreover, intra-individual and inter-individual variability levels remain stable during the time course of the experiments (Figure 3.6B). In summary, we have shown that the behavior of each larva is stable over days and that the intra-variability is

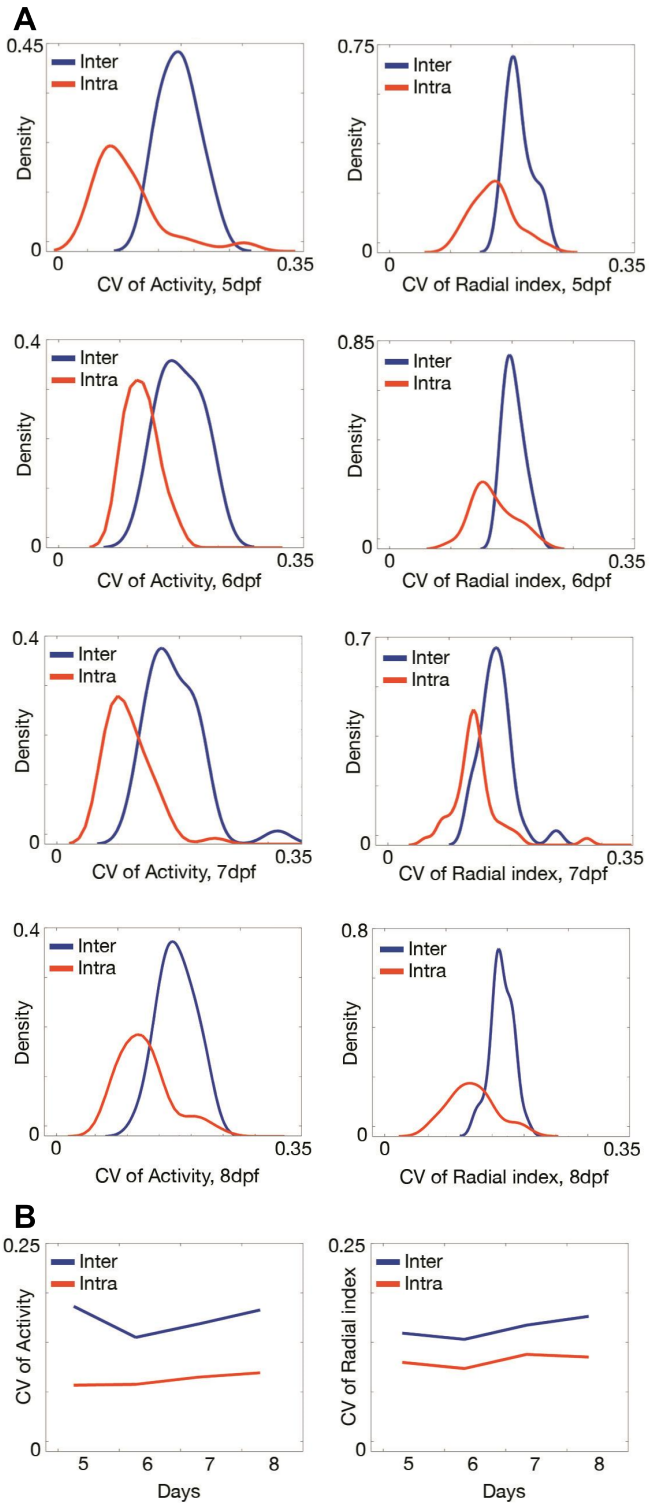

FIGURE 3.6: (A) Smoothed histogram of the coefficient of variation of the activity (left column) and the radial index (right column) showing the intra-variability (red) and inter-individual variability (blue) of the same group at from 5 to 8 dpf. (B) Representation of the median CV of intra-variability (red) and inter-variability (blue) for activity (left) and radial index (right) during the time course of the experiments

smaller than the inter-variability for the parameters we analyzed, so our experimental model for zebrafish larva under uniform environments can be employed to study behavioral individuality.

3.2.2 Sources of behavioral variability in zebrafish

Our next step was to test the relevance of genetic differences and environmental factors for the behavioral variability of a population. In order to find whether two animal groups had different behavioral variability, we decided to compare the generalized variance of both groups. To compute the P-value we generated 1,000 random realizations of the two groups. Each realization was obtained by mixing all individuals from the two groups and randomly extracting two new groups. Then, we computed the P-value as the proportion of random repetitions whose value of the difference of generalized variance was equal or higher than the experimental one.

We first analyzed whether uncontrolled environmental changes underlie behavioral variability. We had already shown that the behavior of larvae did not depend specifically on their position on the plate, but our experimental conditions also minimized other possible environmental variations: eggs were isolated in multiwell plates at pharyngula stage (24 hours post-fertilization), when the brain primordium is still in an immature stage (Kimmel et al., 1995), the temperature was controlled and the changes in water and feeding were carried out in the same conditions for every larva. Nevertheless, the latter are manual processes, and minor changes across the larvae could lead to behavioral alterations between individuals. Thus, we assessed if the removal of these manual techniques could alter the behavioral variability of a population. We compared the behavioral variability of a group of fish with and without a water change 24 hours before the experiment, and found no difference (Figure 3.7.A, top, P=0.42). We then recorded simultaneously two groups of 7 dpf fish that had been raised being or not being fed, and we again found no differences in their behavioral variability (Figure 3.7A bottom, P=0.38). Therefore, the behavioral variability seems to be resistant to some degree of environmental changes (Figure 3.7.A′), even though we cannot control other non-shared influences such as in which position the larva was when its well was moved for experimental reasons. Other factors like parental effects transported in the eggs (Andersson and Höglund, 2012) or even in the sperm Seong et al. (2011) might also affect behavioral variability.

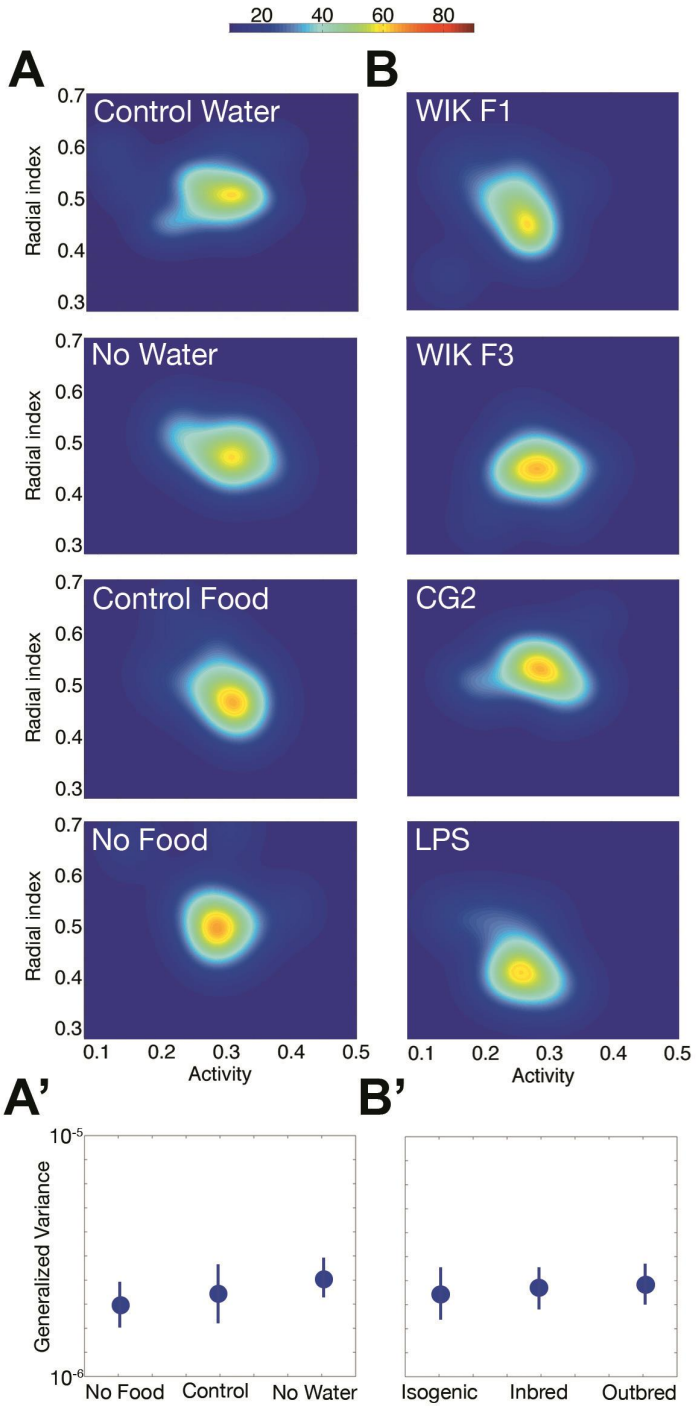

FIGURE 3.7: (A) Probability density of finding an individual with a given mean activity and radial index for groups with water changes and without water changes, with control food or without food at 7 dpf. (A′) Quantification of the generalized variance of each group. (B) Probability density of finding an individual with a given mean activity and radial index for 7 dpf groups with different genetic backgrounds: WIK F1 (three cycles of inbreeding), WIK F3 (two additional inbreeding cycles), CG2 (gymnogenetic fish clones) and LPS (outbred parents). (B′) Quantification of the generalized variance of each group.

We then studied the dependence of the behavioral variability on the genetic differences of the individuals. We used populations with different genetic backgrounds to check if these differences were reflected in the behavioral variability. Our control inbred WIK zebrafish population (F1) was the result of a single batch of eggs retrieved from two adults with at least three cycles of inbreeding. We generated two more inbreeding cycles between siblings (populations F2 and F3) trying to further decrease the genetic differences between siblings. However, we could not see a decrease in the inter-individual behavioral variability (WIK F3, Figure 3.7B, P=0.33 and WIK F2, P=0.32). As a control with isogenic conditions, we used a clonal homozygous population of zebrafish that had been obtained by double heat-shock gymnogenesis (Mizgirev and Revskoy, 2010), but again found no difference in variability as compared to the WIK F1 population (CG2, Figure 3.7B, P=0.44). To study a population with genetic variability higher than WIK, we used siblings from genetically diverse outbred parents and we again found no difference in behavioral variability (LPS line, Figure 3.7B, P=0.38). These results show that behavioral variability is largely independent of genetic differences in the population (Figure 3.7B′, see section 3.6.2 for further details of the zebrafish lines) and takes place even under uniform environments (Figure 3.7A′).

3.2.3 Chromatin acetylation and deacetylation alter behavioral variability

We next assessed the importance of epigenetic factors in the generation of behavioral variability, because they could act as a coding mechanism of the factors that may affect individual behavior. One of the main epigenetic marks, histone acetylation, is related to changes in the gene expression, and is involved in several neuronal processes (Grunstein, 1997; Valor et al., 2013). We used a class I Histone Deacetylase (HDAC) inhibitor, sodium butyrate (NaBu), to determine the contribution of this epigenetic mark to behavioral individuality. Interestingly, sodium butyrate added to the water not only provoked a change in the behavioral variability, but also did not increase it as we could have expected if the role of HDACs was to maintain reduced levels of variability. Instead, it severely reduced the behavioral variability of a WIK F3 sibling population 24h after the treatment (2 mM NaBu, Figure 3.8.A) compared to control PBS-treated larvae (PBS, Figure 3.8.A, P<0.001). Note that this treatment only altered variability and not the mean of the population parameters (P=0.63). When we retired the treatment, behavioral

variability was recovered after additional 24h (Figure 3.8.B, P=0.71). To control that sodium butyrate had the expected effect of HDAC inhibition at the molecular level, we confirmed that there was an increase in the total acetyl-Histone 4 (acH4) amount in larval protein extracts. Similarly to the behavior, this increase was lost after retiring the treatment (Figure 3.8.C). To confirm the role of HDAC inhibition on behavioral variability, we used an hdac1-mutant heterozygotic population (Noël et al., 2008) that also showed a reduced behavioral variability (Figure 3.8.D, P=0.008).

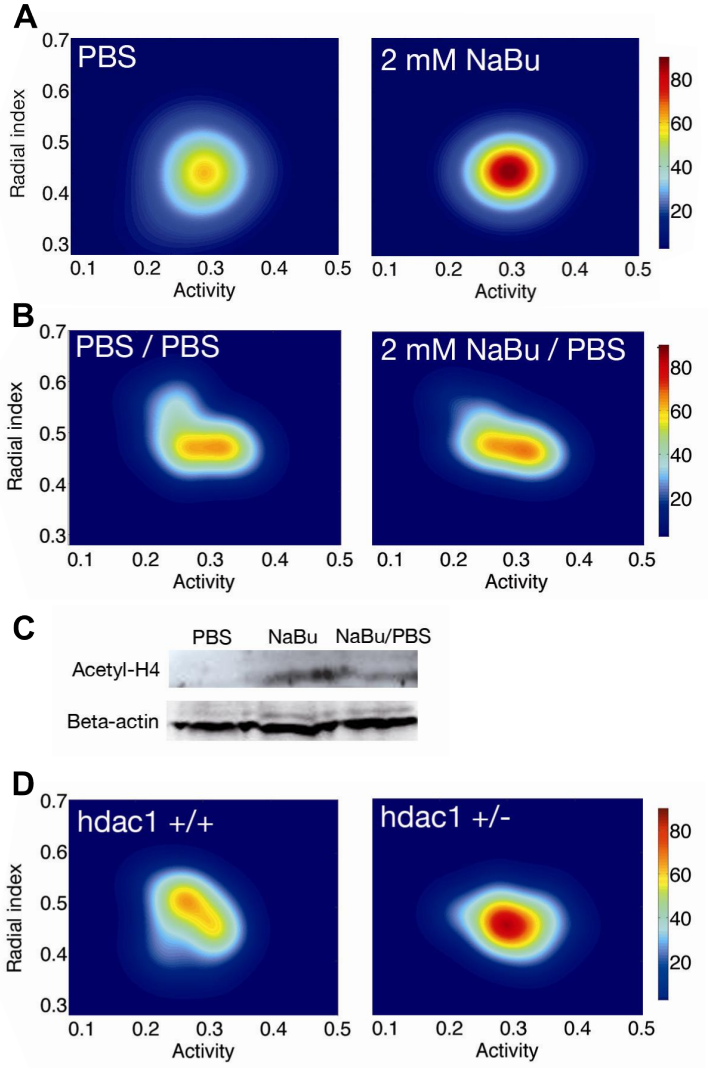

FIGURE 3.8: (A) Probability density of finding an individual with a given mean activity and radial index for fish treated with a PBS solution as control and sodium butyrate as HDAC inhibitor. (B) Same probability as in A for larvae 48h after treatment with sodium butyrate only during the first 24h and then washed with PBS (PBS/PBS as control). (C) Western blot (bottom) showing the amount of acH4 in fish extracts after PBS, NaBu and NaBu/PBS treatments, using -actin as loading control. (D) Same probability as in Figure 3A for hdac1 +/+ and hdac1 +/- larvae.

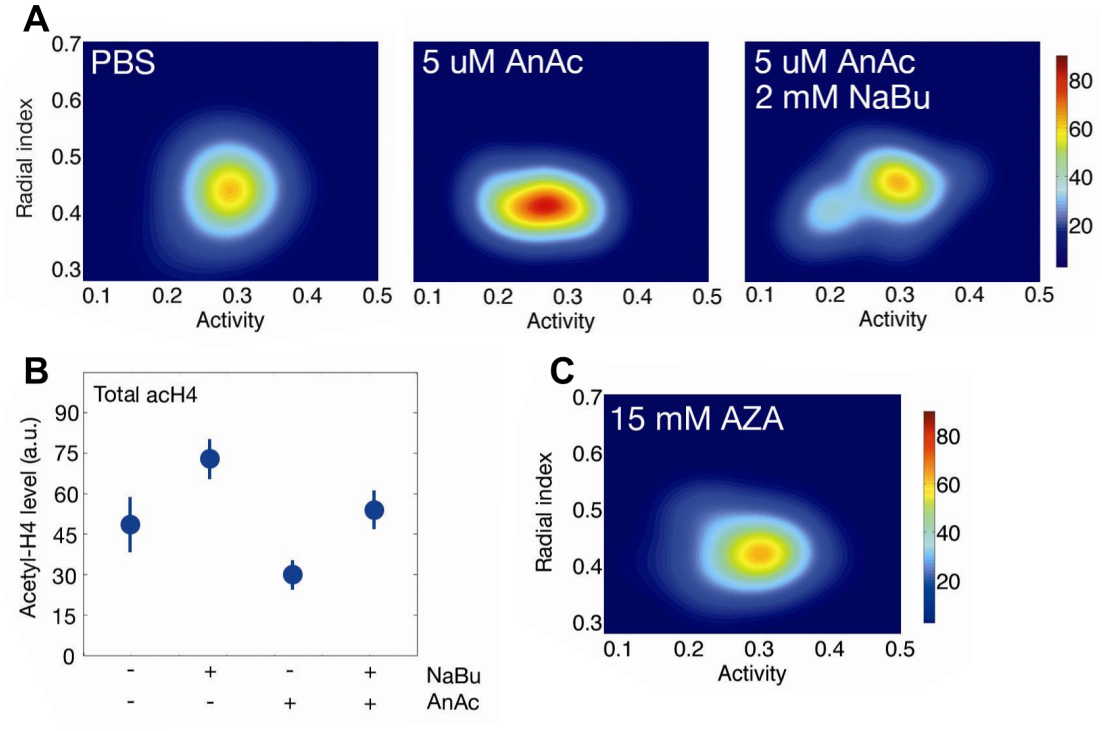

FIGURE 3.9: (A) Probability density of finding an individual with a given mean activity and radial index for fish treated with a PBS solution as control, anacardic acid as HAT inhibitor and double-treated animals, with sodium butyrate and anacardic acid. (B) Total acH4 levels in NaBu, AnAc and double treatments. (C) Same probability as in A for larvae 24h after treatment with 15 mM AZA.

After assessing the role of HDACs in the generation of behavioral variability, we turned to analyze the process of histone acetylation. When we used anacardic acid (AnAc) as a general inhibitor of Histone Acetyl-Transferases (HATs), we obtained results similar to NaBu-treated larvae (5 μM AnAc, Figure 3.9A, P=0.009). This was unexpected as histone acetylation and deacetylation are opposite processes, and AnAc produced a decrease in total acH4 levels in the animals, the reversed response to the NaBu treatment (Figure 3.9B). A possible explanation would be that the balance between HDAC and HAT activity is needed for the control levels of behavioral variability. Consistent with this possibility we found that a double treatment of NaBu and AnAc reverted larval behavioral variability (Figure 3.9A, P=0.005) and total acH4 (Figure 3.9B) to control levels. Other epigenetic alterations, like the DNA-methyltransferases inhibition by 5-azacytidine (AZA) did not alter the behavioral variability of the zebrafish, suggesting the specificity of the histone acetylation pathway (15 mM AZA, Figure 3.9C, P=0.44).

The result that histone acetylation has a large impact on behavioral individuality led us to explore the epigenomic regions that could be the targets of the histone

acetylation pathway. Specifically, we searched for regions with the highest acetylation variability as candidate substrates of behavioral variability. For that, we built four clusters of sibling fish with low intra-cluster behavioral variability and high inter-cluster variability, so that they represented the complete phenotypic space of the population. We analyzed our population in clusters for two reasons: (1) the amount of tissue in a single larva was not sufficient to measure histone acetylation and (2) the clustering reduced the noise that could arise from the molecular analysis of a single larva. To define the clusters, we used a Hierarchical Clustering analysis using Euclidean distance as the metric and the average linkage clustering as the linkage criteria. We chose the four clusters that minimized the distance between its elements and, at the same time, maximized the distance between clusters (clusters c1-c4 in Figure 3.10A and its representation on the phenotypic space in Figure 3.10B).

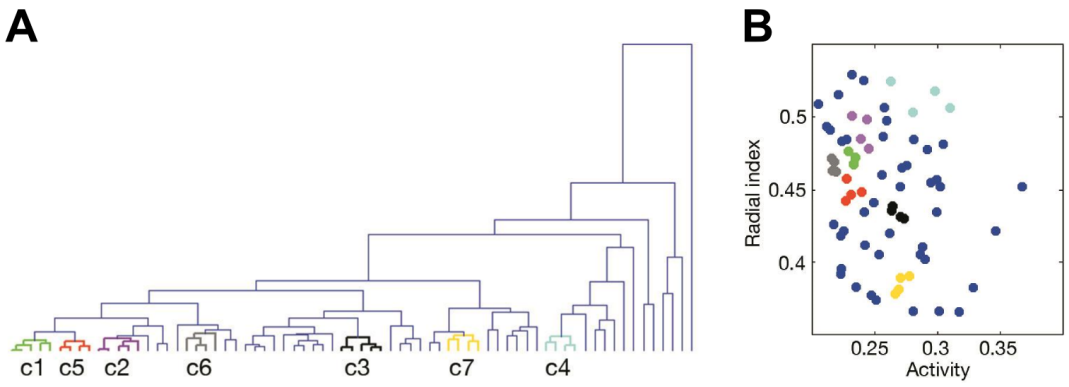

FIGURE 3.10: (A) Dendrogram indicating the distances between clusters, as obtained after the Hierarchical Clustering Analysis. Clusters chosen for epigenomic map are indicated by a color and a specific name. (B) Average activity-radial index plot of the same fish, showing again the differences among groups. Dark blue points indicate not chosen animals.

Considering each cluster as a representative sample of a specific behavioral profile, we retrieved their acH4 epigenomic profiles using chIP-seq. Then, we analyzed the genomic regions with the highest variability in histone acetylation between the different clusters of fish (Figure 3.11A, $P \leq 0.05$, see also **refAppendixA**). Biological ontology of genes located near the detected hyper-variable acH4 regions gave a significant enrichment in terms involved in neuronal differentiation, cell adhesion and motion, axonogenesis and axon guidance (Figure 3.11B, $P=0.004$). Using conventional chIP for three additional fish clusters (c5-c7 in Figure S3F), we confirmed the acetylation hypervariability in eight selected regions (blue genes in Figure 3.11C, top and center line, $P<0.01$) compared to control regions with

low variability (red genes). In addition, the mRNA levels of the genes located near these regions also reflected a similar high-variability status between clusters (Figure 3.11C, bottom line, P<0.01), showing that the hypervariable acetylation of close regions probably lead to the hypervariable expression of these genes.

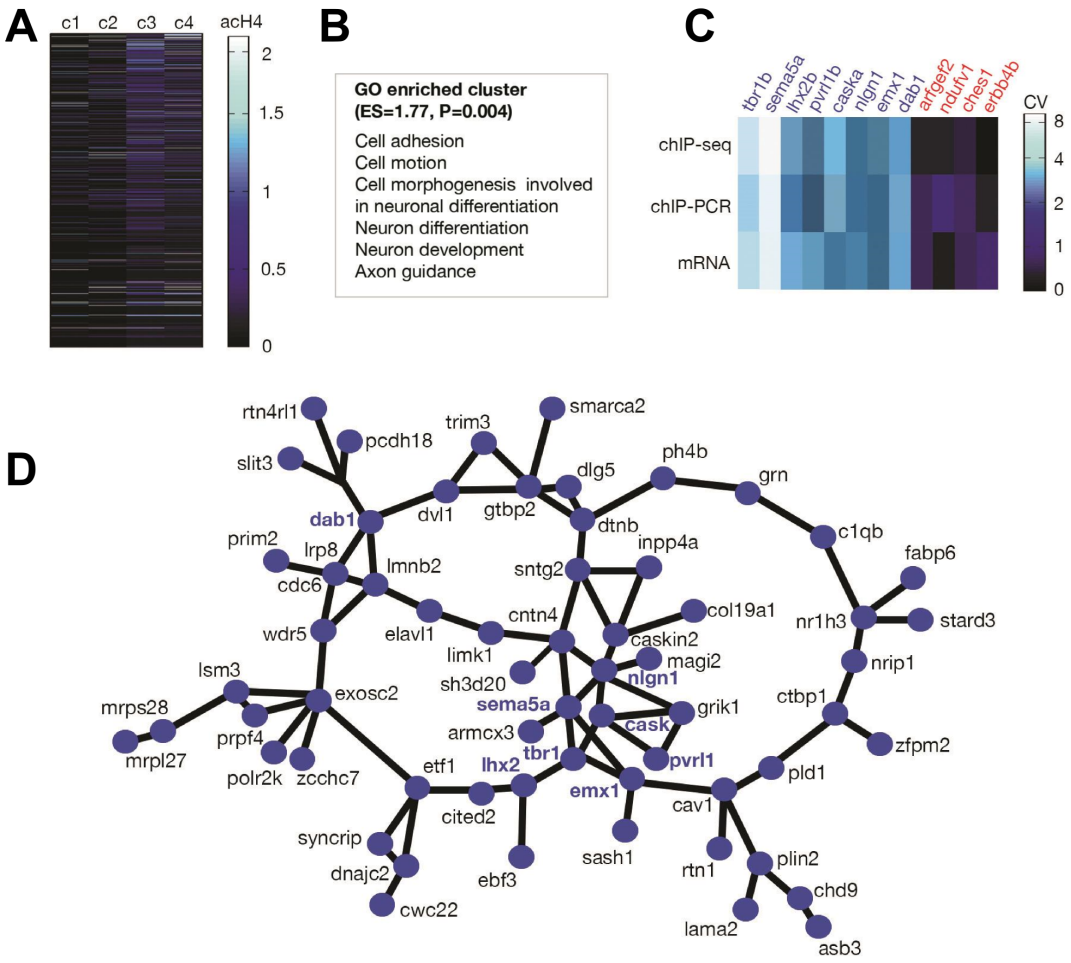

FIGURE 3.11: (A) AcH4 epigenome map in four different groups of 4 sibling fish of low behavioral intravariability and high behavioral intervariability (c1, c2, c3 and c4). Regions are ordered from top to bottom by their acetylation variability across the c1-c4 groups. (B) Results from the GO analysis of the genes located near hypervariable acetylated regions. (C) Coefficient of variation of several hyper-variable (in blue) and low-variable regions (in red) using chIP-seq experiment, using conventional chIP with other behavioral groups (c5, c6 and c7) and also using the mRNA level of their neighboring genes. (D) Predicted network of functional associations between the human orthologs of the hyper-variable genes

We used predicted associations between human orthologs of the genes obtained to generate a functional network, which we will refer to as hypervariable network (Figure 3.11D, right). A core module in this network is composed of a

group of genes that maintain many internal associations between them (edge density=0.174), and integrate the rest of the members (edge density=0.026). Members of this core module are related to neural development, and include proteins involved in synaptic and axon development (SEMA5A, NLGN1, PVRL1), transcriptional modulators (EMX1, LHX2), and members of the reelin signaling pathway (TBR1/CASK transcriptional complex, and DAB1) among others. In summary, this functional network is a promising candidate for the generation of individual differences in behavioral parameters.

We then decided to study the molecular response of these genes to HDAC or HAT inhibition, using the same treatments as before. In this case, to analyze the data we chose 4 clusters of fish as representatives of the behavioral variability of each larval population (see Experimental Procedures, section 3.6.12, for details). We found that the variability of the clusters in histone acetylation (Figure 3.12A) and mRNA expression (Figure 3.12B), measured as the Coefficient of Variation (CV), was altered after the treatments. Both HDAC and HAT inhibition reduced the CV of both histone acetylation and mRNA expression of the hypervariable genes (Figure Figure 3.12A and Figure 3.12B, blue genes) compared to control genes (Figure 3.12A and 3.12B, red genes; $P<0.001$ for all cases) while the double treatment produced a reversion of the CVs to control level (Figure 3.12A and 3.12B, $P=0.18$ and $P=0.25$, respectively). This implies that the response of the genes selected from the hypervariable network is concomitant with the behavior of the animals (Figure 3.8A and Figure 3.9A).

The decrease of the molecular variability in single NaBu and AnAc treatments was accompanied by an increase in the histone acetylation and gene expression in both cases (Figure 3.13A and 3.13B, $P=0.004$ and $P=0.005$, respectively) compared to control regions. In the case of AnAc treatment, this result was unexpected because the global response of HAT inhibition provoked a reduced acetylation (Figure 3.9B), so it confirmed the strong dependency between HAT and HDAC in the individuality network regions. In addition, the double treatment reverted acH4 and mRNA to control levels (Figure 3.13A and 3.13B, $P=0.45$ and $P=0.66$, respectively). Our results showed a decrease of behavioral, epigenetic and transcriptional variability after HDAC or HAT inhibition due to impaired levels of acetyl-H4 in the hypervariable network regions. In this way, these hypervariable genes are suggested to be the final effectors of the pathway required for behavioral variability through different histone acetylation patterns. This relation is supported by the verification that the expression of two transcription factors of

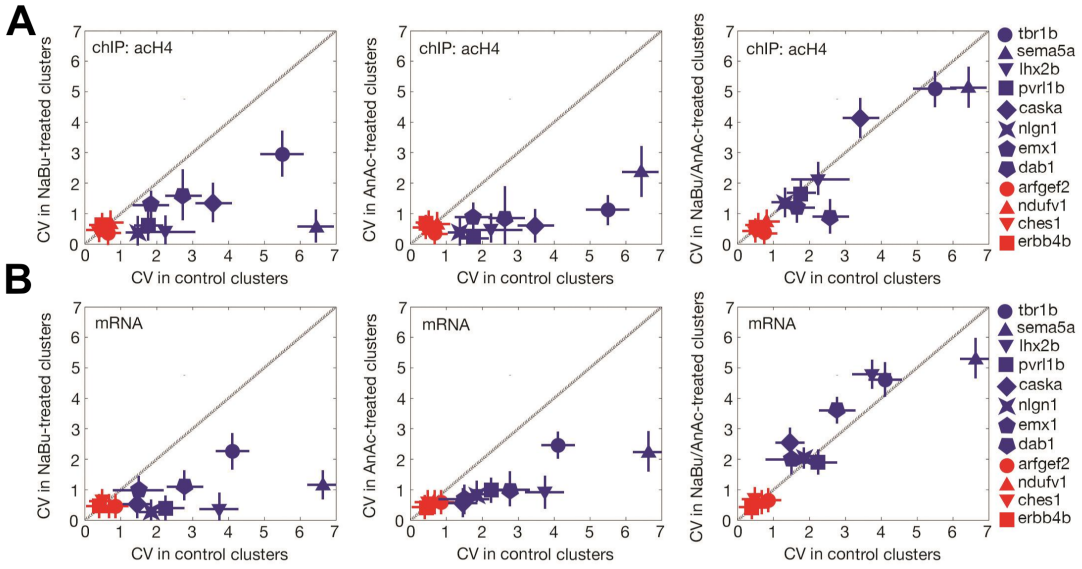

FIGURE 3.12: (A) Coefficient of Variation of acetyl H4 regions after NaBu, AnAc and NaBu/AnAc treatment (from left to right) plotted against the Coefficient of Variation of the same regions in control clusters. Blue dots represent hyper-variable regions and red dots low-variable regions. The correspondence between the symbols and the genes is described in the legend on the right. (B) Same as A, but with the Coefficient of Variation from mRNA gene expression instead.

the core module, extracted from single larvae, correlated with their behavioral parameters (Tbr1b correlated with activity, see 3.14A; P=0.031; Lhx2b correlated with radial index, see 3.14B P<0.001).

3.2.4 Yin-Yang 1 drives molecular and behavioral variability

So far we have shown that behavioral variability requires an epigenetic substrate which modulates a hypervariable gene network related to neural development. This epigenetic variability could be driven by transcriptional factors that guide the HDAC/HAT complex to target regions. Candidates were retrieved by searching for enriched DNA motifs near the acH4 hyper-variable sequences of the network. Several DNA motifs (P<0.0001) were significantly located in these regions (Figure 3.15), including sequences similar to known binding sites for Yin-Yang 1 (YY1), PAX6, LHX and Forkhead-Box families. One of these transcription factors, YY1 is a GLI-Krppel zinc-finger family transcription factor conserved in metazoa that is involved in many developmental and differentiation processes (Donohoe et al., 1999; He et al., 2010). It can activate or repress the same target gene depending

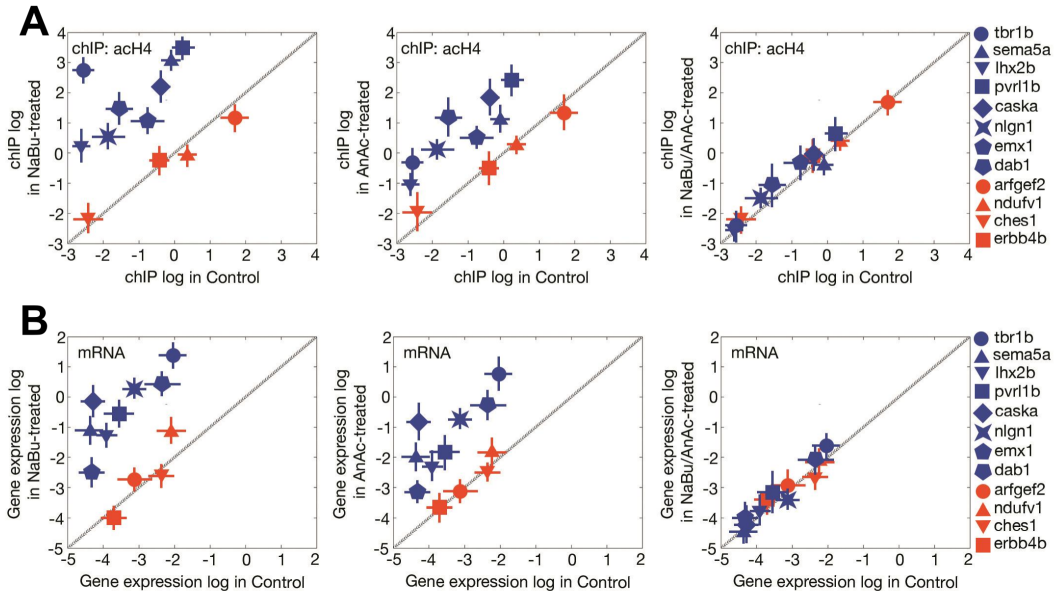

FIGURE 3.13: (A) acH4 fold change in hypervariable (blue) and control (red) regions after NaBu, AnAc and double treatments, respectively. Symbols are described in the legend on the right. (B) Same as A, but using mRNA gene expression fold change.

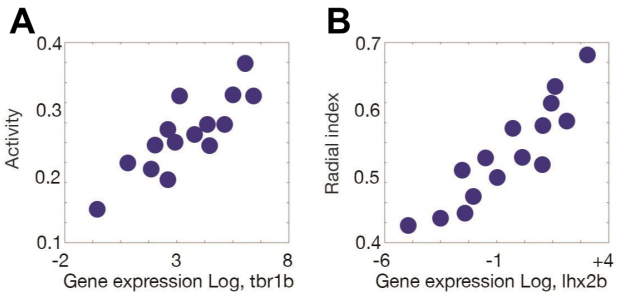

FIGURE 3.14: Correlation between gene expression and behavioral parameters for individual fish. (A) Activity correlated tbr1b expression. (B) Radial index correlated with lhx2b expression.

on recruited co-factors and cell type (He et al., 2010; Shi et al., 1991); in addition, HDACs and HATs (Last et al., 1999; Yao et al., 2001) are among its known cofactors, making YY1 a candidate for the regulation of behavioral individuality.

Therefore, we assessed the binding of YY1 to the selected regions from the hypervariable network, and its modulation by the drugs used previously. In PBS-treated larvae used as control, YY1 was found to be bound to several regions of the hypervariable network with a near YY1 binding site (Figure 3.16A, $P<0.001$). Furthermore, this interaction was impaired when we treated the animals with either NaBu or AnAc (Figure 3.16A, $P<0.001$), and returned to control levels with the double treatment (Figure 3.16A, $P=0.51$). Interestingly, the levels of YY1

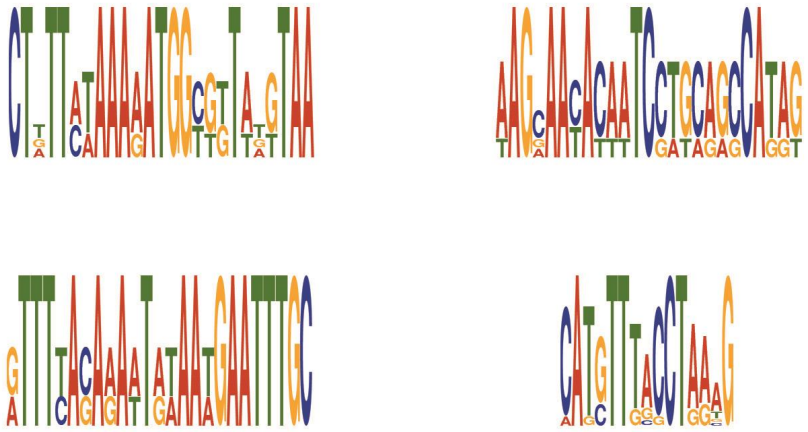

FIGURE 3.15: DNA motifs enriched in the hyper-variable acetylated regions, with similar sequences to YY1 (top left), PAX6 (top right), LHX (bottom left) and Forkhead-Box (bottom right) binding sites.

binding correlated with the behavioral and molecular variability we had shown in Figures 3.8A, 3.9A and 3.12: in larvae treated with NaBu or AnAc there was a concomitant reduction in both binding of YY1 and behavioral and molecular variability, compared to PBS and double treatments.

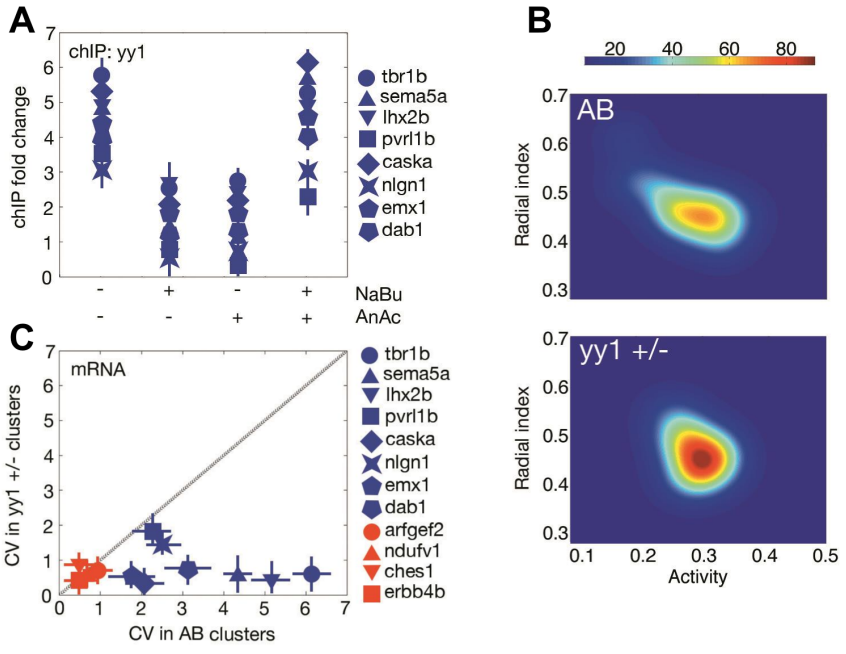

FIGURE 3.16: (A) YY1 chIP binding to hypervariable regions of control, NaBu, AnAc and double treatments shown as fold change compared to the average of low variability (arfgef1, ndufv1, ches1 and erbb4b) regions. Symbol names are described in the legend. (B) Same probability as in Figure 3.7 for AB siblings and yy1 +/- larvae. (C) Coefficient of Variation of mRNA gene expression in behavioral clusters of yy1 +/- larvae compared to AB fish. Symbols are described in the legend.

These results suggest a direct participation of YY1 in the pathway required for behavioral variability in zebrafish. In fact, we found that YY1 heterozygotic mutant fish, generated by in vitro fertilization of AB eggs with YY1 mutant sperm (Wang et al., 2007), had less behavioral variability than the AB control population (Figure 3.16B, P=0.003). When we used clusters of control AB and YY1 mutant fish, we observed a lower variability in the mRNA levels of the selected hypervariable genes (Figure 3.16C, P<0.001).

Interestingly, the reduction in the molecular variability of YY1 heterozygotic fish was acquired without significant changes in the average acetylation and mRNA of the individuality network genes (Figure 3.17A, P=0.64 and P=0.27, respectively). This is an important difference from the results obtained using inhibitors of HDAC and HAT, as these treatments increased average acetylation and mRNA levels in the gene network (Figure 3.13A and 3.13B). This result led us to analyze the role of the inhibitors in the YY1 heterozygotic fish.

We found that there was no further decrease of behavioral variability in YY1 mutants treated with NaBu (Figure 3.17B, P=0.34) and that the expression levels of genes of the hypervariable network were similar to the PBS-treated fish (Figure 3.17C, P=0.44). Unexpectedly, inhibition of HATs by AnAc treatment turned out to be lethal for YY1 heterozygotic larvae, as it is the case of YY1 homozygotic mutant fish. These results point out that a shared pathway involving histone acetylation enzymes and YY1 is necessary for the generation of behavioral individuality.

3.2.5 Acetyl-CoA levels regulate behavioral variability

We have shown that YY1, HDACs and HATs are part of a pathway required for behavioral individuality in zebrafish larvae. YY1 can be regulated by lysine acetylation (Yao et al., 2001) and the molecular activity of HDACs and HATs is related to this signaling pathway. This prompted us to study the role of lysine acetylation in behavioral variability. The main substrate for these reactions is acetyl-CoA, a key metabolite for several pathways like the citric acid cycle and the lipid metabolism. Recently, it has been described how the acetyl-CoA levels control cell growth and proliferation in yeast through the modulation of histone acetylation in specific genes (Takahashi et al., 2006). We thus decided to analyze if the alteration of acetyl-CoA levels in fish could influence their behavioral individuality,

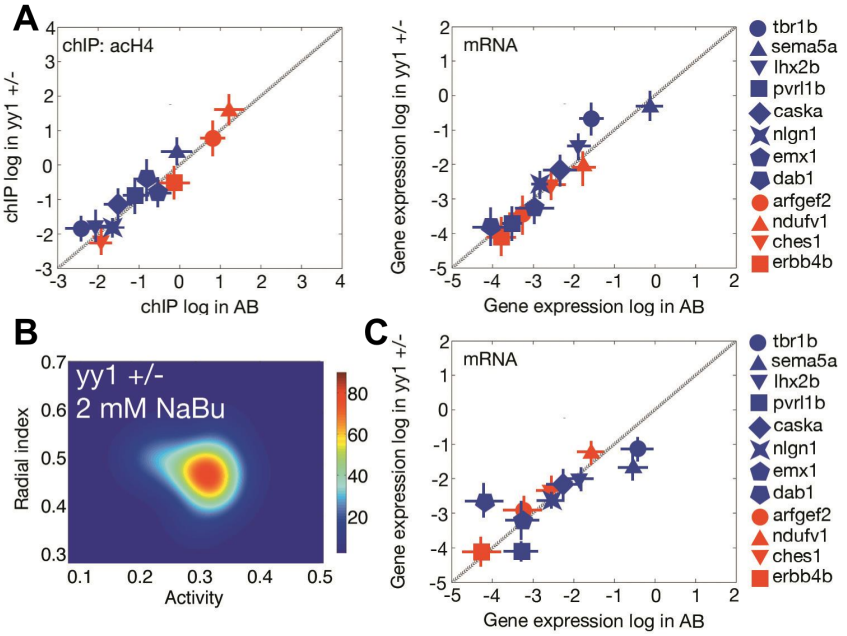

FIGURE 3.17: (A) acH4 (left) and mRNA (right) fold change in hypervariable (blue) and control (red) regions in yy1 +/- animals compared to AB larvae. Symbols are described in the legend on the right. (B) Same probability as in Figure 3.16B for yy1 +/- larvae after treatment with sodium butyrate (left). (C) Same as A but for mRNA fold change after NaBu treatment of yy1 +/- larvae.

by using drugs that alter acetyl-CoA metabolism: Hydroxycitric acid (HCA) as an antagonist of the ATP citrate lyase; MEDICA-16 (MED-16), as an inhibitor of the acetyl-CoA carboxylase; and cerulenin (CER), which can antagonize with the Fatty acid synthase. When we analyzed the behavior of fish treated with these drugs for 24 hours compared to DMSO-treated larvae as vehicle control, we observed a decrease in behavioral variability of HCA-treated larvae (Figure 3.18A, P=0.001), and an increase in the behavioral variability of MED-16- and CER-treated animals (Figure 3.18A, P=0.003 and P=0.009). Furthermore, HCA produced a decrease of the acetyl-CoA and total acH4 levels compared to control, and both MED-16 and CER treatments resulted in an increase of the acetyl-CoA and acH4 levels (Figure 3.18B).

The regulation of acetyl-CoA levels by HCA and MED-16 are direct consequences of their action decreasing its generation or degradation, respectively. However, the action of CER might be indirect, mediated through an excess of malonyl-CoA which eventually results in high levels of acetyl-CoA, possibly through an increased malonyl-decarboxylase activity. Acetyl-CoA levels are potently regulated by metabolic state, but we observed no difference in terms of behavioral variability between normally fed and food deprived larvae (Figure 3.7) and we also

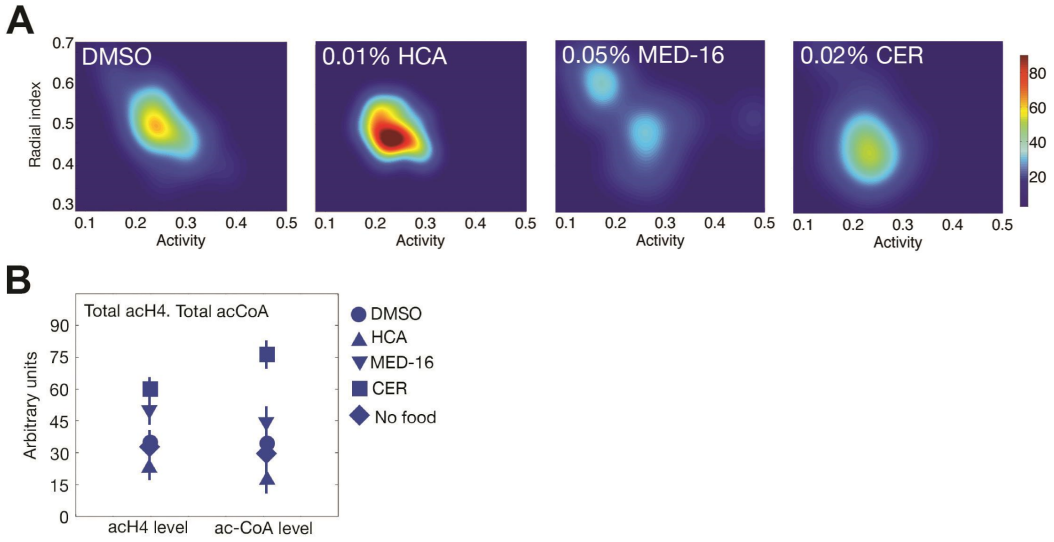

FIGURE 3.18: (A) Probability density of finding an individual with a given mean activity and radial index for control (DMSO), 0.01% hydroxycitric acid (HCA), 0.05% Medica-16 (MED-16) and 0.02% cerulenin (CER) siblings. (B) Total acH4 and acetyl-CoA levels in different pools of larvae. Symbols are described in the legend.

confirmed that the 7 dpf larvae deprived of food had also similar acetyl-CoA and acH4 levels as compared to control larvae (Figure 3.18B). Then we compared the variability of the mRNA expression of the selected genes from the hypervariable network with that of the control genes (Figure 3.19). We observed a reduction of the CV in the HCA-treated fish (Figure 3.19, $P<0.001$), and an increase of the CV in both MED-16 and CER-treated animals (Figure 3.19, $P<0.001$ for both cases). This was similar to the observed changes in behavior, and correlated with the acetyl-CoA levels in the population.

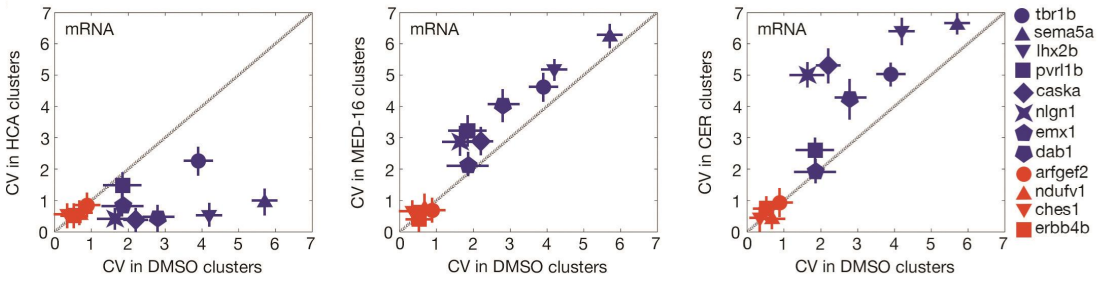

FIGURE 3.19: Coefficient of Variation of mRNA gene expression in behavioral clusters after the treatment with HCA, MED-16 and CER, respectively. Symbols are described in the legend on the right.

We analyzed the acetyl-CoA levels of clusters of control fish with low intra-cluster and high inter-cluster behavioral variability, and found no direct correlation between either activity or radial index with the acetyl-CoA state of these animals (Figure 3.20A). Intriguingly, high levels of acetyl-CoA were present in clusters of fish with activity values far from the average. This could suggest a mechanism by which acetyl-CoA influences the behavioral variability of the population, with higher levels corresponding to extreme behavioral parameters. In order to analyze the role of acetyl-CoA within our pathway, we studied the YY1 binding to the hypervariable regions after modulation of acetyl-CoA levels. YY1 is depleted from these regions in HCA-treated larvae (Figure 3.20B, $P<0.001$) similar to NaBu or AnAc-treated animals, while YY1 binding is increased compared to control after the treatments with MED-16 or CER (Figure 3.20B, $P=0.02$ and $P=0.004$, respectively). This supports the idea that YY1 binding in the selected hypervariable regions is a proxy of behavioral variability, and acetyl-CoA levels could act as an upstream regulator of histone acetylation. To test this hypothesis, we co-treated the fish with MED-16 and NaBu to observe the acetyl-CoA effect in absence of HDAC activity. Our results (Figure 3.20C, $P=0.005$) showed that the behavioral variability was similar to the one produced by NaBu alone, and that acetyl-CoA depended on a functional HDAC pathway to affect this pathway.

3.2.6 Conservation of the hypervariable pathway in humans

We finally tested whether this molecular substrate found in developing zebrafish could in principle be maintained in other species. Using a public gene expression dataset of human pre-natal brains (Colantuoni et al., 2011), we analyzed the expression of conserved YY1 target genes, and we found they also had hyper-variable expression compared to targets of other transcription factors (Jiang et al., 2007) or random genes (Figures 3.21A, and 3.21B, $P=0.001$). In addition, the hypervariable network considered globally had higher variability than random networks (Figure 3.21B, in red, $P=0.002$). Interestingly, some of the genes with hypervariable expression in human brains are TBR1 and LHX2 (Figure 3.21C, $P<0.05$) whose expression correlated with the behavioral parameters analyzed in fish (Figure 3.14). Moreover, human genes with hypervariable expression are significantly enriched in Gene Ontology terms like neuron differentiation, axonogenesis or axon guidance, similar to the zebrafish genes (Figure 3.11B).

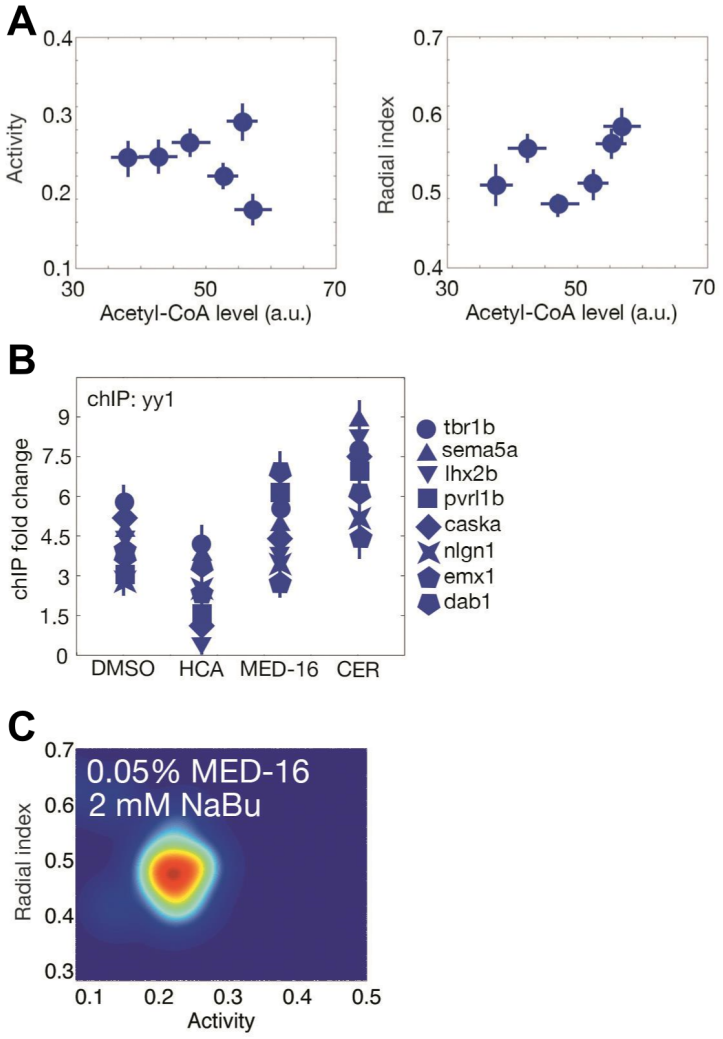

FIGURE 3.20: (A) Acetyl-CoA levels in behavioral clusters of zebrafish, compared to the average activity (left) and the average radial index (right) of these clusters. (B) YY1 chIP binding to hypervariable regions of DMSO, HCA, MED-16 and CER treatments shown as fold change compared to the average of low variability (arfgef1, ndufv1, ches1 and erbb4b) regions. Symbol names are described in the legend. (C) Same probability as in Figure 3.18A, but using a double treatment of NaBu and MED-16.

For a further confirmation, we analyzed the expression of the genes in the hypervariable network using the Connectivity Map (CMAP) drug profiling database (Lamb et al., 2006), that includes the expression profiling in drug-treated human cells. We obtained the drugs which significantly altered the expression of the components of the hypervariable network in each cell line ($P<0.01$, 3.21D) and, interestingly, the most affected pathway by these drugs is the HDAC activity, even in different cell lines (numbers in parenthesis in Figure 3.21D). In addition, other drugs are related with acetyl-CoA (apigenin is a Fatty Acid Synthase inhibitor like cerulenin), and other pathways like PI3K, HSP90, EGFR or D2R. These findings

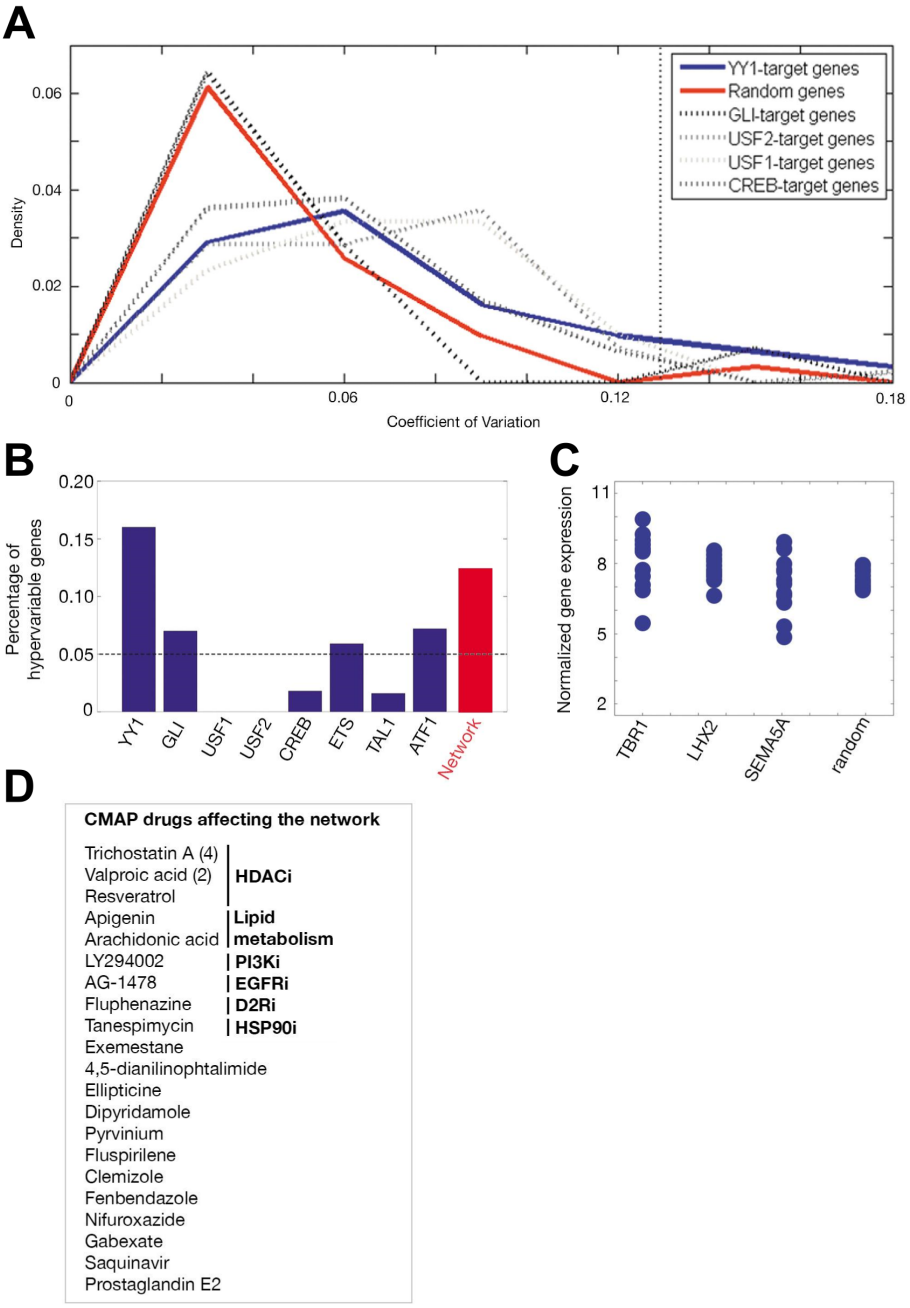

FIGURE 3.21: (A) Comparison of the distributions of transcriptional variability (using pre-natal human brains) of random genes (red), YY1-target genes (blue) and targets from other transcription factors (gray, dashed). (B) Percentage of hypervariable ($P<0.05$) target genes of several transcription factors and the hypervariable network using pre-natal brain gene expression dataset. Dotted line indicates the percentage of genes for the random case. (C) Examples of the individual expression of some genes from the network and random ones in the human pre-natal brains. (D) Full list of CMAP drugs that alter the expression of a significant number of human orthologous genes from the hypervariable network. Some relevant biological pathways affected by the drugs are highlighted beside their names.

show a possible conservation of the pathway in molecular variability in humans.

3.3 Discussion

Our results show that the behavioral individuality requires the balance of YY1, HDAC and HAT activities, which might be regulated by acetyl-CoA levels. The interaction of these proteins generates different levels of histone acetylation which result in individual mRNA profiles of a hypervariable gene network involved in several molecular roles like neuronal migration, development and formation of mature neural circuits (Figure 3.22).

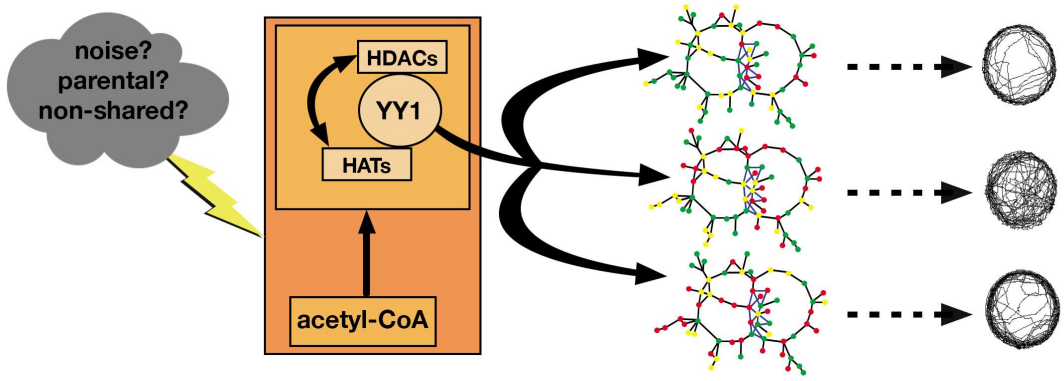

FIGURE 3.22: Schematic model of the action of acetyl-CoA, YY1, HDACs and HATs in the generation of individual epigenetic, transcriptional and behavioral profiles of genetic-independent fish, with possible upstream factors.

We therefore propose a molecular pathway involving acetylation that is required for behavioral individuality in a vertebrate species. This link found between acetylation and behavioral individuality opens many avenues for future research. Our results were independent of the genetic differences among individuals and obtained in identical environments under uniform conditions.

The specific contribution of the different sources that can generate variability in these conditions, such as developmental noise, parental effects or the different experiences individuals may get by interacting with the environment, might be addressed in further studies. Our results suggest that these sources would be translated, through the action of a complex mechanism involving HDAC, HAT, YY1 and acetyl-CoA, into different individual histone acetylation and mRNA patterns, which finally might modulate behavioral variability. Thus, future studies could try to analyze the impact that the different sources of variability may have on our mechanism.

The molecular mechanism by which YY1, in a complex with HATs and HDACs, would produce different outputs (histone acetylation profiles) in each fish should also be addressed in future studies, maybe involving transcription factor binding probabilities (Grönlund et al., 2013) or transcriptional noise (Little et al., 2013). YY1 stands for Yin-Yang 1 because it activates or represses the same target genes depending on different factors like HDACs or HATs (He et al., 2010; Last et al., 1999; Shi et al., 1991; Yao et al., 2001). The results of this paper allow us to propose a mechanism by which YY1 drives variability in acH4 and mRNA profiles of hypervariable genes: (i) YY1 heterozygotic animals were shown to be insensitive to NaBu treatment, suggesting that HDACs require a functional YY1 protein to exert their activity on hypervariable genes, although we cannot discard that a minimum behavioral variability was reached in the YY1 mutants; (ii) HATs activity seems necessary for YY1 to become active, as HAT inhibition was lethal for YY1 heterozygotic mutants and the phenotypes for AnAc and NaBu treatments were very similar; (iii) the molecular and behavioral recovery of double (NaBu and Anac)-treated animals is consistent with a possible necessary balance between HDAC and HAT activities controlled through YY1. At this point, we cannot discard the direct modulation of YY1 activity through its acetylation or deacetylation by the same HDACs and HATs (Yao et al., 2001). Our results propose a novel but logically expected function for YY1 in the generation of molecular variability in individuals. We cannot discard the participation of other proteins in this process, as we found enrichment of other transcription factor binding sites like PAX6, LHX or Forkhead-Box families.

The role of specific members of the hypervariable network should be further analyzed, but the final effect on behavior might be a consequence of the interactions between several hypervariable genes. It would be interesting to analyze the role of hypervariable genes, like TBR1 and CASK, which are known to regulate the expression of reelin and other genes. Reelin is a crucial player in synaptic plasticity processes (Herz and Chen, 2006), which might become a target of the network. The additional presence in this network of LRP8, a reelin receptor, and DAB1, a key regulator of reelin signaling, whose mutation (Scrambler) mimics the one of reelin (Reeler) (Sheldon et al., 1997), supports this hypothesis. Even though we cannot discard the participation of other organs besides the brain in the behavioral variability, the modulation of synaptic plasticity through the hypervariable network might mediate changes in behavior. In this way, the role of acetyl-CoA as a metabolite influencing the level of behavioral variability seems to be important, as

it is also a key sensor of the metabolic state of the animal. A possible explanation is that acetyl-CoA levels regulate the activity of the YY1/HATs/HDACs complex. Alternatively, other molecular substrates might be downstream of acetyl-CoA.

Our results suggest that a molecular program generates behavioral variability in larval zebrafish without the influence of genetic differences. Instead of a control mechanism repressing these alterations, this hypothesis would point to a beneficial effect of the behavioral variability for the zebrafish larvae, similarly to what has been observed in other cellular and physiological processes (Chang et al., 2008; Wernet et al., 2006). Then, this program could be a feature of species with high offspring rate and low survival rate to confront rapidly changing environments; nevertheless, our preliminary results in human brains seem to point to a similar overlapping program in mammals, with low offspring rate and high survival rate. Consequently, other species might be able to adapt to novel environments by producing variable non-heritable traits in their offspring. Finally, the modulation of this or other similar pathways could help us to understand and control the variability in several diseases or drug-response profiles, where histone acetylation has been precisely shown to regulate stochastic heterogeneity within the patient (Sharma et al., 2010).

3.4 Conclusions

- Larval zebrafish present behavioral variability independent of genetic differences and environmental factors. We designed an experimental model to study individuality in behavior by showing that the activity and the radial index of each larva is stable for days and independent of its position on the setup.

- Histone acetylation is directly related to the generation of behavioral invididuality. Treatments that modify the activity of Histone Deacetylases or Histone Acetyl-Transferases modify the behavioral variability of a population.

- Variability in acetylation of different genomic regions predicts the existence of a gene network that could mediate the existence of behavioral variability. The variability was also confirmed with the mRNA expression of these genes, even relating two of them with the studied behavioral parameters.

- The transcription factor Yin-Yang 1 drives molecular and behavioral variability in our studied regions. The binding of Yin-Yang 1 correlates in all cases with the variability of the population.

- Acetyl-CoA levels are capable of regulating the molecular variability of a population. The use of treatments that modify the levels of Acetyl-CoA proved that these levels also correlate with the behavioral variability of a population.

3.5 Conclusiones

- Las larvas de pez cebra presentan una variabilidad comportamental independiente de diferencias genéticas o factores medioambientales. Hemos diseñado un modelo experimental para estudiar individualidad comportamental mostrando que la actividad y la distancia media al centro del pocillo de cada larva es estable a lo largo de los días e independiente de su posición en el setup.

- La acetilación de histonas está directamente relacionada con la generación de individualidad comportamental. El uso de tratamientos que modifican la actividad de las enzimas Histona Deacetilasa o Histona acetil-transferasa modifican la variabilidad comportamental en una población.

- La variabilidad en la acetilación de distintas regiones genómicas predice la existencia de una red de genees que podría mediar la variabilidad comportamental. La variabilidad también se confirmó con la expresión por mRNA de estos genes e incluso se relacionó a dos de ellos con los parámetros comportamentales analizados en este estudio.

- El factor de transcripción Yin-Yang 1 es capaz de dirigir la variabilidad comportamental y molecular en las regiones que hemos estudiado. La unión de Yin-Yang 1 correlaciona en todos los casos con la variabilidad en la población.

- Los niveles de Acetyl-CoA son capaces de regular la variabilidad molecular de una población. El uso de tratamientos que modifican los niveles de Acetyl-CoA permitieron además comprobar que estos niveles correlacionan con la variabilidad comportamental de la población.

3.6 Materials and Methods

3.6.1 Ethics statement

All the experiments using animals were approved and performed following the guidelines of the CSIC (Spain) for animal bioethics.

3.6.2 Zebrafish lines and care

Zebrafish (Danio rerio) WIK strain (Nechiporuk et al., 1999) was kindly provided by Dr. Bovolenta (CBM-UAM) and inbred in our laboratory for at least three generations before the experiments. Afterwards, WIK F1 population was generated from a single batch of embryos from a couple of adult fish. Two additional cycles of inbreeding (F2 and F3) were carried out, crossing a couple of siblings from the former generation. CG2 clone population, generated by double gymnogenetic heat-shock, and characterized by being pure isogenic zebrafish was kindly provided by Dr. Revskoy (Univ Northwestern) as a control of reduced genetic differences between siblings. The outbred LPS (Local Pet Store) strain was recently described (see Arganda et al. (2012)), and used as a model of genetic heterogeneity. Heterozygotic hdac1 mutant strain with wild-type counterparts was a kind gift by Dr. Ober (NIMR). Heterozygotic Yin-Yang 1 (YY1) mutant zebrafish generated by in vitro fertilization of YY1 mutant sperm of AB eggs were obtained, as AB eggs, from ZIRC zebrafish repository. Both homozygotic hdac1 and YY1 mutations become lethal for zebrafish.

Parent adult fish were kept in the animal facility with a 14/10 light/dark cycle, in 5-l or 8-l transparent containers connected to a larger fish rack system with circulating water at 26.5 $\pm$ 0.5 C and at 77.5 pH. Fish densities were up to 1.25 fish per liter. Fish were fed live artemia (Artemia salina) twice a day, and fish flakes (Sera Vipal) once a day. Water conditions were maintained using appropriate filters and were measured once a week in order to keep low levels of $NH3$, $NO2$ and $NO3$. In the case of our experimental larvae, eggs were isolated after 24 hours post-fertilization (hpf), and maintained in custom multiwell plates until 10 days post-fertilization (dpf). They were fed (JBL NovoBaby) and water-changed daily from 4 dpf if it is not indicated specifically in the experiment.

3.6.3 Free-swimming setup and recording

The camara used was a 1.2 MPixel monochrome camera (Basler A622f, with a Pentax objective of focal length 16 mm) and was placed at a distance of 70 cm over the wells and pointing downwards. The wells are 15 mm deep, and have a diameter of 20 mm at the bottom and a diameter of 32.5 mm at the top. The dishes , with 24 wells each, are supported by a white PMMA surface that is only partially opaque. Behind this white surface we place two infrared led arrays (830nm, TSHG8400 Vishay Semiconductors) pointing outwards. Two paper sheets stand between the lights and the central space that lies directly under the wells. With this disposition we ensure that only diffuse indirect light reaches the wells, so that the illumination is roughly uniform (most of the light comes from below the wells through the white surface). All the set-up is surrounded by white curtains. Video camera recorded at a 30 fps rate. A larval population (5-8 dpf) consisted of at least 24 fish siblings from the same batch of embryos. After five minutes of acclimation to the new environment, the larvae were recorded for 20 minutes. Water temperature was maintained in a strict range (25.5-27°C) during each experiment.

3.6.4 Data Analysis and Statistical Analysis

Several MatLab scripts were used to do the mathematical analyses of the paper. Tests using randomly generated data to compute a P-value used 1,000 permutations. All the experimental procedures were done at least three times with different biological datasets, and P-values were calculated using the three replicas. Figures show a representative experiment of the triplicate.

3.6.5 Gaussian smoothing algorithm for representing the variability of a population

A simple visual method to characterize the variability in a population is to plot the bi-dimensional distribution of the mean activity and radial index of individuals (Figures 3.5). We used a nice visualization of this distribution using Gaussian kernel smoothing that consists in adding up Gaussians centered at the data points as

$$P(x,y) = \frac{1}{N}\frac{1}{2\pi\sigma_x\sigma_y}\sum_{i=1}^{N} exp\left\{-\frac{1}{2}\left[\frac{(x-x_i)^2}{\sigma_x^2} + \frac{(y-y_i)^2}{\sigma_y^2}\right]\right\} \quad (3.1)$$

with x_i and y_i are the mean activity and radial index values of individual i of a total of N members of the population. An optimal smoothing uses standard deviations of each Gaussian given by $\sigma_x = N^{-\frac{1}{6}}\alpha_x$ with α_x the standard deviation in the x_i data values, and similar for σ_y using the y_i values (see B.E. Hansen, unpublished manuscript, http://www.ssc.wisc.edu/ bhansen/718/NonParametrics1.pdf).

3.6.6 Comparison of the significance in variability using different parameters

We ran statistical tests using other parameters to compare the variability between different populations. We chose as parameters the Standard Deviation (SD) and the Coefficient of Variation (CV) for both behavioral parameters (activity and radial index) separately. The results are similar to those obtained with the *Generalized Variance* and can be seen in the following table.

Experiment	SD P-value (activity)	SD P-value (radial index)	CV P-value (activity)	CV P-value (radial index)
WIK F1 vs F3	0.73	0.13	0.43	0.24
WIK F1 vs CG2	0.43	0.10	0.04	0.88
WIK F1 vs LPS	0.79	0.70	0.91	0.76
Control Food vs No Food	0.44	0.71	0.54	0.75
Control Water vs No Water	0.67	0.72	0.88	0.81
PBS vs NaBu	<0.001	<0.001	<0.001	<0.001
PBS vs Anacardic acid	0.07	0.012	0.003	0.002
PBS vs AnacA + NaBu	0.78	0.85	0.45	0.88
PBS vs NaBu (48h)	0.73	0.77	0.71	0.61
Hdac1 +/+ vs hdac1 +/-	0.54	<0.001	0.68	<0.001
AB vs YY1 +/-	0.02	<0.001	0.011	<0.001
YY1 +/- vs YY1 +/- NaBu	0.41	0.28	0.36	0.30
DMSO vs HCA	<0.001	<0.001	<0.001	<0.001
DMSO vs MED-16	0.004	<0.001	0.002	<0.001
DMSO vs CER	0.009	0.001	0.007	<0.001
DMSO vs MED-16 + NaBu	0.001	0.008	<0.001	0.004

3.6.7 Reagents and antibodies

Sodium butyrate (*Sigma-Aldrich*) was dissolved in Phosphate-buffered saline (PBS), and used in a final 2 mM concentration of fish water. PBS alone was used as vehicle control. Ethanol solution of anacardic acid (*Sigma-Aldrich*) mixed with fish water (final concentration: 5 μM) was tested, compared with ethanol alone, and there were no evident changes to the PBS control. 5-Azacitidine (*Sigma-Aldrich*) was used with a final concentration of 15 mM. While hydroxycitric acid (*Sigma-Aldrich*; final concentration: 0.01% w/v) was water soluble, cerulenin (*Sigma-Aldrich*; final concentration: 0.02% w/v) and Medica-16 (*Sigma-Aldrich*; final concentration: 0.05% w/v) were dissolved in DMSO. The same final volume of DMSO was added to the water control and to the hydroxycitric acid solution. Acetyl-Histone 4, YY1 and β-Actin antibodies were obtained from *Promega*, *SantaCruz* and *Sigma-Aldrich*, respectively.

3.6.8 Western Immunoblotting

3.6.8.1 Sample preparations

Treated and untreated groups of fish (5-10) were frozen at different times, and then protein extracts were isolated from tissue using an extraction buffer (80 mMTris-HCl pH 7.5, 2 mM EDTA, 2 mM EGTA, 0.27 M Saccharose, 10 mM β-glycerolphosphate, 5 mM Sodium pyrophosphate, 50 mM Sodium Fluoride, 1% Triton X-100, 0.1 mM Sodium vanadate, 0.1% β-Mercaptoethanol, 1X Complete protease inhibitor cocktail) during 30 minutes at 4°C with vortexing. Samples were centrifuged at 10000g for 10 min at 4°C. After centrifugation, debris was removed and protein extracts were stored at -20°C for further use, if not used directly.

3.6.8.2 Measurement of protein concentration

The concentration of protein was measured using Bradford protein assay, with the *Coomassie plus protein assay reagent (Pierce)* reagent and bovine serum albumin (BSA) as standard. After the colorimetric reaction, the absorbance of the samples was measured at a wavelength of 595 nm, and the values obtained were linearly interpolated to a calibration curve built with known amounts of BSA.

3.6.8.3 Protein electrophoresis and transference to nitrocellulose membranes

In order to perform the electrophoresis of proteins in polyacrylamide gels, we used the discontinuous buffer system described by Laemmli (LAEMMLI, 1970). Aliquotes containing between 10 and 30 μg of the corresponding protein extracts were mixed with the adequate volume of loading buffer for proteins (5 x loading buffer for proteins: Tris-HCl 32.2 mM pH 6.8, SDS 10% (w/v), glycerol 50% (v/v), bromophenol blue 0.025% (w/v), β-ME 20% (v/v)) and denaturalized by incubation at 100°C for 5 minutes. Proteins were separated according to their molecular weight by electrophoresis in polyacrylamide gels under denaturing conditions, in the presence of SDS (Electrophoresis buffer: Tris 25mM, glycine 190 mM, SDS 0.1% (w/v)). The final concentration of polyacrylamide gel was 10

3.6.8.4 Protein transfer to nitrocellulose membrane

After the electrophoresis, proteins were transferred from the SDS gel onto nitrocellulose membranes (*Bio-Rad Laboratories*) by applying an electric field in a specific transfer buffer (Tris 25mM, glycine 190mM, methanol 20% (v/v)). The transfer was performed at 4 °C and 50 V for 2h. Then, membranes were quickly rinsed with distilled water and protein bands detected with a Ponceau-S stain (Ponceau S 0.5% (w/v), acetic acid 5% (v/v)). After washing with water, we blocked the membranes for 2 h at room temperature in a blocking buffer (1X TBS with 5% w/v nonfat dry milk) and incubated with primary antibody (acH4 antibody with a 1:1000 dilution and load control β-actin, diluted 1:1500). Then, the membranes were washed four times with TBS-T (Tris-HCl mM pH 7.5, NaCl 75 mM, Tween-20 0.2% (v/v)) by agitating for 15 minutes each time, and incubated for 45 minutes with the corresponding secondary antibody conjugated to Horseradish Peroxidase enzyme (HRP). The membranes were again washed three times with TBS-T buffer for 15 minutes and one additional time for 5 minutes in TBS (Tris-HCl 50mM pH 7.5, NaCl 75mM). Finally, the protein bands were detected using chemiluminescent substrate *SuperSignal Chemiluminescent Substrate (Pierce)* for 5 minutes at room temperature, and visualized on KodaK BioMax light film (*Sigma-Aldrich*).

3.6.9 Histone 4 acetylation assay and acetyl-CoA fluorometric assay

Groups of 5 fish were frozen, homogenized using an automatic dounce, and then processed following manufacturer (*Epigentek*) recommendations. In the case of acetyl-CoA quantification, a similar approach was used, following specific manufacturer (*Abcam*) guidelines.

3.6.10 Chromatin Immunoprecipitation (ChIP)

For chIP experiments, clusters of five fish with low intra-cluster behavioral variation and high inter-cluster behavioral variation were frozen, crosslinked with 1.8% formaldehyde for 30 minutes and then quenched with 1% glycine for 5 minutes. Extracts were lysed using a SDS Lysis buffer (50 mM Tris-HCl pH 8.1, 1% SDS, 10 mM EDTA) for 30 minutes at 4°C, and then diluted with a Dilution buffer (6.7 mM Tris-HCl pH 8.1, 0.01% SDS, 1.2 mM EDTA, 1.1% Triton X-100, 167 mMNaCl). 2 mM sodium butyrate was added to avoid histone deacetylation activity during the preparation. Then, the fish were sonicated with two pulses (30 seconds ON / 30 seconds OFF) of 15 minutes each with the Diagenode Bioruptor. DNA fragments were centrifuged at 4°C for 10 minutes at 13000 x G. At this point, an input DNA sample was obtained. To reduce non-specific immunoglobulin binding, samples were pre-cleared for 45 minutes at 4°C in presence of Protein A/G Beads. Then, the extracts were immunoprecipitated overnight at 4°C using 1 μg of the anti- acetyl-Histone 4 antibody. Bound DNA was recovered with an additional incubation with protein A/G beads for 1 hour at 4°C and samples were then centrifuged at 2000 x G for 1 minute at 4°C. After this process, they were washed with Low-Salt (120 mM Tris-HCl pH 8, 0.1% SDS, 2 mM EDTA, 1% Triton X-100, 150 mM NaCl), High-Salt (120 mM Tris-HCl pH8, 0.1% SDS, 2 mM EDTA, 1% Triton X-100, 500 mM NaCl), LiCl (10 mM Tris-HCl pH 10, 1 mM EDTA, 0.25 M LiCl, 1% NP40, 1% sodium deoxycholate) and two times with 1X TE (10 mM Tris-HCl pH 8, 1 mM EDTA) buffers, and recovered with Elution (1% SDS, 0.1 M $NaHCO_3$, for 15 minutes). DNA purified samples were de-crosslinked using sodium chloride (200 mM for 4 hours at 65°C), and cleared with *Qiagen* spin columns. qPCR was used to calculate the percentage of recovered DNA respect to input of each sample, as a measure of the acetylation of H4 or YY1 in specific regions, and p-values obtained by Student's T-test. In the case of YY1, an

additional normalization was done to measure chIP fold change in hypervariable regions against low variable regions.

3.6.11 ChIP-seq analysis

ChIP-seq with anti- acH4 antibody was performed as normal chIP, but final samples (around 2 ng) were processed at the Genomics Unit at the Scientific Park of Madrid. Libraries were built, and the samples were sequenced using an Illumina GAII. Raw data of the experiment can be obtained in the NCBI GEO repository (GSE IDXXX). Reads were aligned to Danio rerio genome sequence (Zv7) with BWA, and final reads in a 25-bp window were mapped to the reference genome using custom Perl scripts. The repetitive regions were removed from the analysis and the results were also normalized using the 90th percentil of the number of reads in each sample. **Numeros de millones de lecturas?** For each sample, we built the statistics for the number of lectures, grouping them depending on the percentile of the histogram they fell (bins=0, 50, 75, 90, 95, 100 percentiles for each sample). Then, we calculated the probability that each region was on a given bin of the histogram normalized to the total number of samples, p_i. For example, if the region was in a different bin for the 4 samples, it would give a 0.25 probability for each sample. However, if all were in the same bin, the probability would be 1. With this probability, we calculated Shannon's entropy $-(\sum_i p_i log(p_i))$ for each region. The final parameter that estimated the variability for each region was the mean of its Shannon's entropy and its CV. Permuted datasets allowed us to estimate the distribution of random variability. Regions with $P<0.05$ in terms of variability were classified as hyper-variable acetylated.

3.6.12 Cluster analysis for conventional ChIP and Acetyl-CoA levels

Our larval population in all the cases apart from the ChIP sequencing consisted of 24 larvae and it was too small to perform a Hierarchical Cluster Analysis, as we did with the ChIP sequencing. We then designed a simpler clustering algorithm for these cases: First, the variability of the population was assumed to be represented by four clusters. Second, the phenotypic space was divided in four quadrants and a clustering algorithm was used for each quadrant, by minimizing the Euclidean

between its elements. Third, the cluster selected for each quadrant was the one that has its mean value near the mean value of all the elements in the quadrants.

3.6.13 RNA isolation and qPCR quantification

3.6.14 Gene ontology

The position of these regions respect to genes was retrieved, and Gene Ontology analysis was performed using DAVID. **DAVID is an algorithm that groups genes based on functional similarity. It clasifies them depending on their similarity based on shared functional annotation using over 75,000 terms from 14 functional annotation sources. In other words, it gives a P-value for the probability that the genes you use as input are grouped in the same functional cluster.** Thus, a cluster of enriched Gene Ontology terms was obtained, and a functional network from human ortholog genes was modeled using STRING (Franceschini et al., 2013). Edge density of sub-networks was calculated as

$$ED(V, E) = \frac{|E|}{|V| \cdot (|V| - 1)} \tag{3.2}$$

with V, E the sets of vertices and edges of the sub-network, respectively. $|E|$ represents the number of edges in set E and $|V|$ the number of vertices within the subnetwork. The denominator of the fraction calculates the number of edges in a fully connected sub-network with $|V|$ vertices. Some of the hypervariable regions were tested in subsequent chIP and mRNA analyses.

3.6.15 Prediction of enriched DNA motifs

We used the MEME Suite (Bailey et al., 2009) to find enriched DNA motifs that were present in our hyper-variable acetylated regions including their flanking sequences. We chose motifs that were expected to have one occurrence per sequence, and we did not search for a specific number of motifs. The motifs that we highlight in the paper are those that had a potential biological activity.

3.6.16 Variability in human gene expression datasets

We calculated the CV for each gene using a public gene expression dataset of 36 human pre-natal human brains. Human orthologs of the individuality network, conserved human targets of YY1, and of other transcription factors were obtained, and the distribution of their CVs was compared to the distribution of random gene sets. The percentage of genes with a CV higher than 98% of the total transcriptome ($P<0.02$) was higher for YY1 targets and the individuality network than for the rest of transcription factors and random networks.

3.6.17 Connectivity map

We also used the Connectivity Map drug profiling database to detect drugs which significantly altered the human orthologs of the individuality network in different tissues. For each condition of drug and tissue, we selected the genes whose expression changed most (the 5% of the total). We found the number of individuality genes that were present in each of these subsets, and using random selection of genes for each condition, we calculated the P-value of the presence of individuality genes in each subset of altered expressed genes.

4. General Discussion

In this thesis, I have studied two different aspects of animal collectives: the influence of group size on social decisions and a molecular pathway involved in the generation of behavioral variability. The relation between these two aspects has poorly been studied and they seem to be unconnected. However, there are some features that could act as liaison and I would like to briefly go through them in this discussion.

From a theoretical point of view, even though the group size did not regulate the variability of a population, it could have a global effect on dividing tasks among specialized group members. For example, we are going to assume that the variability of a population was represented by a Gaussian with its mean and standard deviation and that these parameters were independent of the size of the group. In this situation, it would be easier to build specialized consistent subgroups of individuals in larger populations, as an effective subgroup needs a minimum number of individuals to become productive. As task-switching has some cost associated (Goldsby et al., 2012), individuals may stick to its labor, which would in turn generate a more competitive population, thus improving its global success. This may result in a more heterogeneous group, with more specialized individuals. This effect could also have other consequences at group level, such as a reduction in the intra-specific competition between its members.

Heterogeneity of a group has been proved to be a key aspect when discriminating among individuals, specially for offspring or mate recognition (Medvin et al., 1993; Tibbetts and Dale, 2007). Being individually discriminated becomes more difficult in larger groups and, hence, large group size may select for increased individuality to facilitate discrimination (Mathevon et al., 2003; Medvin et al., 1993). In fact, a pioneering study by Pollard and Blumstein (2011) showed a strong positive link between social group size in sciurid rodents and individuality in their social alarm calls. They calculated the amount of individuality contained in vocal alarm

calls and compared this measure across eight phylogenetically controlled species. Their work is also important because signaling or attending to individual identity may be crucial in a wide range of social communicative encounters, such as group decisions.

Individuality is expected to have an impact on collective decision making (Couzin and Krause, 2003). Different motivations or internal states may result in different individual responses. For example, satiated individuals would be expected to respond less strongly when presented with a food stimulus, compared to hungry fish. There is also likely to be variation in the perceived stimuli and in the inherent propensity of individuals to respond to them(Magurran et al., 1995). Thus, the accuracy of a social decision is going to depend completely on the state or on the capacity of the individuals, with a group of more informed or more competent individuals reaching a better precision. In fact, one of the weakest features of the type of models I used in Chapter 2 is that they assume, for the sake of simplicity, that all individuals are identical. Even though there have been some theoretical works studying this aspect (see Couzin et al. (2002); Romey (1996); Viscido et al. (2007)), we are far from a real understanding how individuality could be implemented in collective decision models, so it should be further analyzed.

To have an experimental insight about how individuality can affect group performance, we conducted a preliminary simple task experiment on groups of adult zebrafish (Pérez-Escudero et al., 2014). We trained four groups of four fish to locate a food patch through 12 different trials and studied the order of arrival at the food in each trial (see Figure 4.1A). In each experiment, the four fish of a group entered the behavioral tank (50 50 2 cm3, width length height) through a door in the middle of one side. The food patch was placed at the corners of the opposite side and was always at the same side in all trials for the same group (two groups were trained with food on the left, and the other two with food on the right). We defined a magnitude, called stability score (s), to measure how consistent the order of arrival was across several trials. We found that groups have different values of the stability score, so the order of arriving at the food patch changes significantly between groups. While two groups exhibited a highly stable order of arrival, the other two had an arrival order compatible with the random case (see Figure 4.1B and C).

With this experiment we see that individuality not only creates differences between individuals, but also generates variability between groups. In some groups,

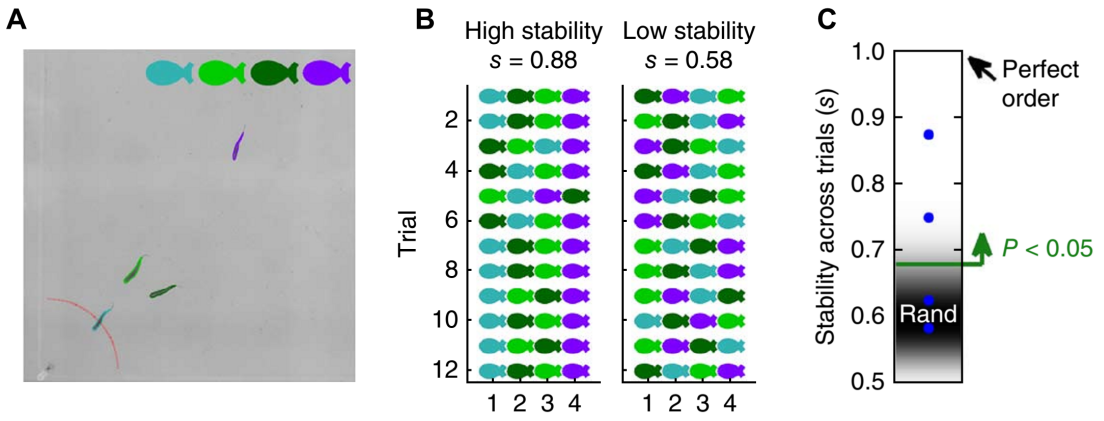

FIGURE 4.1: (A) Frame illustrating one trial of a food-finding task in zebrafish, with colors representing identities. Silhouettes on the top indicate order of arrival. Upon reaching a threshold (red arc), a fish is considered to have arrived at the food. (B) Order of arrival for each trial for two of four total groups, one group with consistent order (left) and another compatible with random order (right). To calculate the value of the stability score, s, we first computed the average rank of each individual across all trials. Then, for each trial and each pair of individuals of the group, we added 1 if the individual with higher rank in the current trial also had higher average rank. We defined s as the resulting number divided by the number of trials and the number of pairs. Thus, s = 1 if all trials have identical ordering. For random ordering, s tends to 0.5 in the limit of a large number of trials. (C) Stability across trials for the four groups (blue dots). The black gradient shows the distribution of s for random ordering (darker for higher probability). To compute the P values, we generated 10,000 random repetitions of the experiment, permuting randomly the order of the individuals in each trial.

the effect of the individualities of their members can create an intrinsic way of organization at an specific task. The first arriving individual may show some specific traits such a higher aggressiveness, dominance or maybe just a better learning ability. Something similar may happen to the last arriving individual, which can be considered as a natural follower of the group. However, we clearly see that other groups do not show this grade of interorganization and this is compatible with individuals being more identical and not presenting a defined role in the group. It may also be the case that the personalities of the individuals in these groups are not so strongly defined, so it is more difficult to see a clear internal structure.

We planned to perform a further experiment on the same task to examine the details of how group composition could condition its performance. In our previous study, fish were placed in groups one day before the experiment, so groups could already have an internal structure at the time the experiment started. This structure could be caused by some specific traits of the individuals and it would be interesting to study their inter-relations before the experiment (see paull2010

for a study in dominance hierarchies in groups of four zebrafish). In our experiments, we may assume that groups with a higher level of internal organization may reach the food in a more stable order. The key point would be to study if this order of arrival is maintained when groups of individuals have not had time to establish an internal structure. In this new experiment, individuals would be kept inside a pool of fish, and would be randomly selected in smaller groups to perform the experiment. If the results of such experiment showed that the groups are highly ordered, it could be assumed that this effect was a result of specific individual characteristics of the individuals and not of the prior internal structure of the groups. One possibility could be that the structure was globally defined in the bigger pool of fish, so it could also be interesting to relate individual behavior in the pool with its performance in the experiment.

The inter-regulation of group behavior by the individualities of its members can also work in the opposite direction. In this sense, being part of a group may modulate the individualities of its members within it too. A recent study by Herbert-Read et al. (2013) tested this hypothesis in groups of fish. They first recorded fish swimming individually and studied how their individuality in movement patterns changed when placed into groups. They showed that the inter-variability and the intravariability were reduced when swimming in group. This suggests that individuals adapt their kinematic characteristics when moving in group. However, behavior was not completely well characterized and the results were not consistent through all the parameters that they studied. Further research should be performed in this field, trying to individually track how each kinematic trait adapts from swimming individually to swimming in group. Another interesting result from the same article is that they found correlation between the loss of individuality and group size, but further research would be needed to more consistently confirm this results.

Even though their work defines a new starting point, it showed some limitations because they restricted the analysis to a specific free swimming task in groups of fish. The regulation of the individuality through group living could act at several levels, and not only creating a conformity between the movement of all the individuals: there may be some molecular mechanisms that are exclusively activated in a social context or the development of some regions of the brain may be modulated by social interactions. A further step to take would be to relate social influence with individuality in a population. It could be the case that group living could act as a buffering mechanism reducing inter-behavioral individuality

whereas individuals raised in social isolation presented higher levels of variability. This is one of the hypotheses that we are trying to develop in the laboratory.

From the human point of view, we could think of situations where it would be very useful to control the variability of a population. For example, in intensive agriculture or cattle industry, it would be more advantageous to control the variability of a very resistant and profitable population. However, this has obviously the implicit risk of reducing population diversity which, in the long run should be ecologically disadvantageous. As a final remark, I like to finish this discussion with a literary comment, by highlighting that some science fiction books, such as *Brave new world* by Aldous Huxley or *1984* by George Orwell, present a distressed and heartbreaking human society where individuality is highly punished. This works may be taken as a warning sign to make us aware of the fact that Social Control should be cautiously applied in our modern society.

Bibliography

Amé, J.-M., Halloy, J., Rivault, C., Detrain, C., and Deneubourg, J. L. (2006). Collegial decision making based on social amplification leads to optimal group formation. *Proceedings of the National Academy of Sciences of the United States of America*, 103(15):5835–5840.

Andersson, M. A. b. and Höglund, E. (2012). Linking Personality to Larval Energy Reserves in Rainbow Trout (Oncorhynchus mykiss). *PLoS ONE*, 7(11).

Arganda, S., Pérez-escudero, A., and Polavieja, G. G. D. (2012). A common rule for decision making in animal collective across species. *Proceedings of the National Academy of Sciences*, 110(9):3651–3651.

Avilés, L. and Tufiño, P. (1998). Colony size and individual fitness in the social spider Anelosimus eximius. *The American naturalist*, 152(3):403–418.

Bailey, T. L., Boden, M., Buske, F. a., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W., and Noble, W. S. (2009). MEME Suite: Tools for motif discovery and searching. *Nucleic Acids Research*, 37(May):202–208.

Berg, S. (1993). Condorcet's jury theorem, dependency among jurors. *Social Choice and Welfare*, 10:87–95.

Bertram, B. C. (1980). Vigilance and group size in ostriches. *Animal Behaviour*, 28(October 1977):278–286.

Biro, D., Sumpter, D. J. T., Meade, J., and Guilford, T. (2006). From Compromise to Leadership in Pigeon Homing. *Current Biology*, 16:2123–2128.

Biro, P. a. and Stamps, J. a. (2008). Are animal personality traits linked to life-history productivity? *Trends in Ecology and Evolution*, 23(May):361–368.

Bode, N. W. F., Faria, J. J., Franks, D. W., Krause, J., and Wood, a. J. (2010). How perceived threat increases synchronization in collectively moving animal groups. *Proceedings. Biological sciences / The Royal Society*, 277(May):3065–3070.

Boinski, S. and Campbell, A. F. (1995). *Use of Trill Vocalizations To Coordinate Troop Movement Among White-Faced Capuchins: a Second Field Test*, volume 132. Brill.

Bolnick, D. I., Amarasekare, P., Ara??jo, M. S., B??rger, R., Levine, J. M., Novak, M., Rudolf, V. H. W., Schreiber, S. J., Urban, M. C., and Vasseur, D. a. (2011). Why intraspecific trait variation matters in community ecology. *Trends in Ecology and Evolution*, 26(4):183–192.

Bolnick, D. I., Svanbäck, R., Fordyce, J. a., Yang, L. H., Davis, J. M., Hulsey, C. D., and Forister, M. L. (2003). The ecology of individuals: incidence and implications of individual specialization. *The American naturalist*, 161(1):1–28.

Bonabeau, E., Dagorn, L., and Fréon, P. (1999). Scaling in animal group-size distributions. *Proceedings of the National Academy of Sciences of the United States of America*, 96(8):4472–4477.

Branco, M., Branco, M., Kidd, N., Kidd, N., Pickard, R., and Pickard, R. (2006). A comparative evaluation of sampling methods for Varroa destructor (Acari: Varroidae) population estimation*. *Apidologie*, 37:452–461.

Brockerhoff, S. E., Hurleyt, J. B., Janssen-bienholdt, U., Neuhauss, S. C. F., Driever, W., and Dowling, J. E. (1995). A behavioral screen for isolating zebrafish mutants with visual system defects. *Proceedings of the National Academy of Sciences of the United States of America*, 92(November):10545–10549.

Brotherton, P. N. M., Russell, A. F., Riain, M. J. O., Gaynor, D., Kansky, R., Griffin, A., Manser, M., Sharpe, L., Mcilrath, G. M., Small, T., Moss, A., and Monfort, S. (2001). Concession in Meerkat Groups. *Science*, 291.

Brown, C. R. (1996). *Coloniality in the Cliff Swallow: The Effect of Group Size on Social Behavior*. University of Chicago Press.

Brown, C. R. and Brown, M. B. (1986). Ectoparasitism as a Cost of Coloniality in Cliff Swallows (Hirundo Pyrrhonota). *Ecology*, 67(5):1206–1218.

Brown, J. L. (1982). Optimal group size in territorial animals. *Journal of Theoretical Biology*, 95:793–810.

Burrill, J. D. and Easter, S. S. (1994). Development of the retinofugal projections in the embryonic and larval zebrafish (Brachydanio rerio). *J Comp Neurol*, 346:583–600.

Chang, H. H., Hemberg, M., Barahona, M., Ingber, D. E., and Huang, S. (2008). Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature*, 453(May):544–547.

Chen, Z. and Kemp, S. (2011). Self-Assessments Produce Anchoring Effects in Promotion Decisions. *European Perspectives on Cognitive Science*.

Clutton-Brock, T. (2002). Breeding together: kin selection and mutualism in cooperative vertebrates. *Science*, 296(5565):69–72.

Clutton-brock, T. (2007). Sexual Selection in Males and Females. *Science*, 318(December):1882–1886.

Cockburn, A. (1998). Evolution of Helping Behavior in Cooperatively Breeding Birds. *Annual Review of Ecology and Systematics*, 29:141–177.

Colantuoni, C., Lipska, B. K., Ye, T., Hyde, T. M., Tao, R., Leek, J. T., Colantuoni, E. a., Elkahloun, A. G., Herman, M. M., Weinberger, D. R., and Kleinman, J. E. (2011). Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature*, 478(7370):519–523.

Condorcet, J.-A.-N. d. C. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*.

Couzin, I., Krause, J., James, R., Ruxton, G., and Franks, N. (2002). Collective memory and spatial sorting in animal groups. *Journal of Theoretical Biology*, 218(1):1–11.

Couzin, I. D. and Krause, J. (2003). Self-Organization and Collective Behavior in Vertebrates. *Advances in the Study of Behavior*, 32:1–75.

Couzin, I. D., Krause, J., Franks, N. R., and Levin, S. a. (2005a). Effective leadership and decision-making in animal groups on the move. *Nature*, 433(7025):513–6.

Couzin, I. D., Krause, J., Franks, N. R., and Levin, S. a. (2005b). Effective leadership and decision-making in animal groups on the move. *Nature*, 433(February):513–516.

da Silva, J. and Terhune, J. (1988). Harbour seal grouping as an anti-predator strategy. *Animal Behaviour*, 36(Mansfield 1967):1309–1316.

Davies, N. B. and Houston, A. I. (1981). Owners and Satellites: The Economics of Territory Defence in the Pied Wagtail, Motacilla alba. *The Journal of Animal Ecology*, 50(1):157.

Dehn, M. M. (1990). Vigilance for predators: detection and dilution effects.

Donohoe, M. E., Zhang, X., McGinnis, L., Biggers, J., Li, E., and Shi, Y. (1999). Targeted disruption of mouse Yin Yang 1 transcription factor results in peri-implantation lethality. *Molecular and cellular biology*, 19(10):7237–7244.

Duckworth, R. a. (2006). Aggressive behaviour affects selection on morphology by influencing settlement patterns in a passerine bird. *Proceedings. Biological sciences / The Royal Society*, 273(April):1789–1795.

Duckworth, R. a. (2008). Adaptive dispersal strategies and the dynamics of a range expansion. *The American naturalist*, 172 Suppl(July 2008):S4–S17.

Duckworth, R. a. (2009). The role of behavior in evolution: A search for mechanism. *Evolutionary Ecology*, 23:513–531.

Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S. (2002). Stochastic gene expression in a single cell. *Science*, 297:1183–1186.

Foster, W. a. and Treherne, J. E. (1981). Evidence for the dilution effect in the selfish herd from fish predation on a marine insect.

Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., Heine-Suñer, D., Cigudosa, J. C., Urioste, M., Benitez, J., Boix-Chornet, M., Sanchez-Aguilera, A., Ling, C., Carlsson, E., Poulsen, P., Vaag, A., Stephan, Z., Spector, T. D., Wu, Y.-Z., Plass, C., and Esteller, M. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences of the United States of America*, 102(30):10604–10609.

Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., Von Mering, C., and Jensen, L. J. (2013). STRING v9.1: Protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Research*, 41(November 2012):808–815.

Franks, N., Bray, H., Hamilton, M., Mallon, E., Mallon, E., Mischler, T., Franks, N., Mallon, E., and Mischler, T. (2003). Strategies for choosing between alternatives with different attributes: Exemplified by house-hunting ants. *Animal Behaviour*, 65:215–223.

Freund, A., Brandmaier, M., Lewejohann, L., Kirste, I., Kritzler, M., Krüger, A., Sachser, N., Lindenberger, U., and Kempermann, G. (2013). Emergence of Individuality in Genetically Identical Mice. *Science*, 340:756–759.

Galef, B. G. and Wigmore, S. W. (1983). Transfer of information concerning distant foods: A laboratory investigation of the information-centre hypothesis. *Animal Behaviour*, 31:748–758.

Galton, F. (1874). *English men of Science: Their Nature and Nurture*.

Galton, F. (1907). Vox populi. *Nature*, 75:450–451.

Gärtner, K. (1990). A third component causing random variability beside environment and genotype. A reason for the limited success of a 30 year long effort to standardize laboratory animals? *Laboratory animals*, 24:71–77.

Gerard, J. F., Bideau, E., Maublanc, M. L., Loisel, P., and Marchal, C. (2002). Herd size in large herbivores: Encoded in the individual or emergent? *Biological Bulletin*, 202(June):275–282.

Giraldeau, L.-A. and Gillis, D. (1985). Optimal group size can be stable: A reply to sibly. *Animal Behaviour*, 33:666–667.

Goldsby, H. J., Dornhaus, a., Kerr, B., and Ofria, C. (2012). Task-switching costs promote the evolution of division of labor and shifts in individuality. *Proceedings of the National Academy of Sciences*, 109(34):13686–13691.

Gomez-Marin, A., Paton, J. J., Kampff, A. R., Costa, R. M., and Mainen, Z. F. (2014). Big behavioral data: psychology, ethology and the foundations of neuroscience. *Nature Neuroscience*, 17(11):1455–1462.

Greene, E. (1987). Individuals in an osprey colony discriminate between high and low quality information. *Nature*, 329:239–241.

Grönlund, A., Lötstedt, P., and Elf, J. (2013). Transcription factor binding kinetics constrain noise suppression via negative feedback. *Nature Communications*, 4(May):1864.

Grunstein, M. (1997). Histone acetylation in chromatin structure and transcription. *Nature*, 389:349–352.

Haas, V. (1985). Colonial and single breeding in fieldfares, Turdus pilaris L.: a comparison of nesting success in early and late broods. *Behavioral Ecology and Sociobiology*, 16:119–124.

He, Y., Kim, J. Y., Dupree, J., Tewari, A., Melendez-Vasquez, C., Svaren, J., and Casaccia, P. (2010). Yy1 as a molecular link between neuregulin and transcriptional modulation of peripheral myelination. *Nature neuroscience*, 13(12):1472–1480.

Heller, E. a., Cates, H. M., Peña, C. J., Sun, H., Shao, N., Feng, J., Golden, S. a., Herman, J. P., Walsh, J. J., Mazei-Robison, M., Ferguson, D., Knight, S., Gerber, M. a., Nievera, C., Han, M.-H., Russo, S. J., Tamminga, C. S., Neve, R. L., Shen, L., Zhang, H. S., Zhang, F., and Nestler, E. J. (2014). Locus-specific epigenetic remodeling controls addiction- and depression-related behaviors. *Nature Neuroscience*, 17(12):1720–1727.

Herbert-Read, J. E., Krause, S., Morrell, L. J., Schaerf, T. M., Krause, J., and Ward, A. J. W. (2013). The role of individuality in collective group movement. *Proceedings of the Royal Society of London B*, 280(1752).

Herz, J. and Chen, Y. (2006). Reelin, lipoprotein receptors and synaptic plasticity. *Nature reviews. Neuroscience*, 7(November):850–859.

Higashi, M. and Yamamura, N. (1993). The University of Chicago. *The American Naturalist*, 142(3):553–563.

Hoare, D. J., Couzin, I. D., Godin, J. G. J., and Krause, J. (2004). Context-dependent group size choice in fish. *Animal Behaviour*, 67:155–164.

Hughes, a. R., Inouye, B. D., Johnson, M. T. J., Underwood, N., and Vellend, M. (2008). Ecological consequences of genetic diversity. *Ecology Letters*, 11:609–623.

Jaenisch, R. and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature genetics*, 33(march):245–254.

Jain, A. K., Prabhakar, S., and Pankanti, S. (2002). On the similarity of identical twin fingerprints. *Pattern Recognition*, 35:2653–2663.

Jiang, C., Xuan, Z., Zhao, F., and Zhang, M. Q. (2007). TRED: A transcriptional regulatory element database, new entries and other development. *Nucleic Acids Research*, 35:140–143.

Kaern, M., Elston, T. C., Blake, W. J., and Collins, J. J. (2005). Stochasticity in gene expression: from theories to phenotypes. *Nature reviews Genetics*, 6:451–464.

Kain, J. S., Stokes, C., and de Bivort, B. L. (2012). Phototactic personality in fruit flies and its suppression by serotonin and white. *Proceedings of the National Academy of Sciences of the United States of America*, 109(48):19834–9.

Kao, A. B., Couzin, I. D., and B, P. R. S. (2014). Decision accuracy in complex environments is often maximized by small group sizes Decision accuracy in complex environments is often maximized by small group sizes. (April).

Kenward, R. E. (1978). Hawks and doves: factors affecting success and selection in goshawk attacks on woodpigeons. *The Journal of Animal Ecology*, 47(2):449–460.

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., and Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Developmental dynamics : an official publication of the American Association of Anatomists*, 203:253–310.

Kimmel, C. B., Patterson, J., and Kimmel, R. O. (1974). The development and behavioral characteristics of the startle response in the zebra fish. *Developmental psychobiology*, 7(1):47–60.

Krause, J. and Ruxton, G. (2002). *Living in Groups*. Oxford series in ecology and evolution. Oxford: Oxford University Press.

Kruuk, H. (1972). *The Spotted Hyena: A Study of Predation and Social Behavior*.

Ladha, K. K. (1992). The Condorcet Jury Theorem, Free Speech, and Correlated Votes. *American Journal of Political Science*, 36(3):617–634.

LAEMMLI, U. K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, 227(5259):680–685.

Lamb, J., Crawford, E. D., Peck, D., Modell, J. W., Blat, I. C., Wrobel, M. J., Lerner, J., Brunet, J.-p., Subramanian, A., Ross, K. N., Reich, M., Hieronymus, H., Wei, G., Armstrong, S. a., Haggarty, S. J., Clemons, P. a., Wei, R., and Carr, S. a. (2006). The Connectivity Map : Using. *Science*, 313(September):1929–1935.

Last, T. J., Van Wijnen, A. J., Birnbaum, M. J., Stein, G. S., and Stein, J. L. (1999). Multiple interactions of the transcription factor YY1 with human histone H4 gene regulatory elements. *Journal of Cellular Biochemistry*, 72:507–516.

Lee, M. D. and Shi, J. (2010). The Accuracy of Small-Group Estimation and the Wisdom of Crowds. *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, pages 1124–1129.

Little, S. C., Tikhonov, M., and Gregor, T. (2013). Precise developmental gene expression arises from globally stochastic transcriptional activity. *Cell*, 154(4):789–800.

Magurran, A., Seghers, H., Shaw, P. W., and Carvalho, G. R. (1995). The Behavioral Diversity and Evolution of Guppy , Poecilia reticdata, Populations in Trinidad. *Advances in the Study of Behavior*, 24:155–2020.

Martinez, F. a. and Marschall, E. a. (1999). A dynamic model of group-size choice in the coral reef fish Dascyllus albisella. *Behavioral Ecology*, 10(5):572–577.

Mathevon, N., Charrier, I., and Jouventin, P. (2003). Potential for individual recognition in acoustic signals: A comparative study of two gulls with different nesting patterns. *Comptes Rendus - Biologies*, 326(3):329–337.

McCann, K. S. (2000). The diversity-stability debate. *Nature*, 405(May).

McElligott, M. B. and O'Malley, D. M. (2005). Prey tracking by larval zebrafish: Axial kinematics and visual control. *Brain, Behavior and Evolution*, 66:177–196.

Medvin, M., Stoddard, P., and Beecher, M. (1993). Signals for parent-offspring recognition: a comparative analysis of the begging calls of cliff swallows and barn swallows. *Animal Behaviour*, 45:841–850.

Milton, K. (2000). Quo vadis? Tactics of food search and group movement in primates and other animals. *On the move (Boinski, S. and Garber, P.A., eds) pp. 357-418, University of Chicago Press.*

Mizgirev, I. V. and Revskoy, S. (2010). A new zebrafish model for experimental leukemia therapy. *Cancer Biology and Therapy*, 9(February 2015):895–903.

Mussweiler, T. and Pfeif, T. (1991). Over coming the Inevitable anchoring Effect : Considering the Opposite Compen sates for Selective Accessibility. pages 1142–1150.

Nabet, B., Leonard, N. E., Couzin, I. D., and Levin, S. a. (2009). Dynamics of decision making in animal group motion. *Journal of Nonlinear Science*, 19:399–435.

Nanjundiah, V. and Bhogle, A. S. (1995). The precision of regulation in Dictyostelium discoideum: implications for cell-type proportioning in the absence of spatial pattern. *Indian journal of biochemistry & biophysics*, 32(6):404–16.

Nechiporuk, A., Finney, J. E., Keating, M. T., and Johnson, S. L. (1999). Assessment of polymorphism in zebrafish mapping strains. *Genome Research*, 9(801):1231–1238.

Nelson, R. J. and Chiavegatto, S. (2001). Molecular basis of aggression. *Trends in Neurosciences*, 24(12):713–719.

Nijhout, H. F., Maini, P. K., Madzvamuse, A., Wathen, A. J., and Sekimura, T. (2003). Pigmentation pattern formation in butterflies: Experiments and models. *Comptes Rendus - Biologies*, 326:717–727.

Nitzan, S. and Paroush, J. (1984). The significance of independent decisions in uncertain dichotomous choice situations. *Theory and Decision*, 17(1984):47–60.

Niwa, H. (1998). School size statistics of fish. *Journal of theoretical biology*, 195(3):351–61.

Niwa, H. S. (2003). Power-law versus exponential distributions of animal group sizes. *Journal of Theoretical Biology*, 224:451–457.

Noël, E. S., Casal-Sueiro, A., Busch-Nentwich, E., Verkade, H., Dong, P. D. S., Stemple, D. L., and Ober, E. A. (2008). Organ-specific requirements for Hdac1 in liver and pancreas formation. *Developmental biology*, 322(2):237–50.

Nowak, M. a. (2006). Five rules for the evolution of cooperation. *Science*, 314(December):1560–1563.

Nowak, M. a., Tarnita, C. E., and Wilson, E. O. (2010). The evolution of eusociality. *Nature*, 466(7310):1057–1062.

Oldroyd, B. P. and Fewell, J. H. (2007). Genetic diversity promotes homeostasis in insect colonies. *Trends in Ecology and Evolution*, 22(8):408–413.

Öst, A., Lempradl, A., Casas, E., Weigert, M., Tiko, T., Deniz, M., Pantano, L., Boenisch, U., Itskov, P. M., Stoeckius, M., Ruf, M., Rajewsky, N., Reuter, G., Iovino, N., Ribeiro, C., Alenius, M., Heyne, S., Vavouri, T., and Pospisilik, J. A. (2014). Paternal Diet Defines Offspring Chromatin State and Intergenerational Obesity. *Cell*.

Packer, C., Scheel, D., and Pusey, A. (1990). Why Lions Form Groups: Food is Not Enough. *The American Naturalist*, 130(1):526–543.

Pérez-Escudero, A. and de Polavieja, G. G. (2011). Collective animal behavior from Bayesian estimation and probability matching. *PLoS computational biology*, 7(11):e1002282.

Pérez-Escudero, A., Vicente-Page, J., Hinz, R. C., Arganda, S., and de Polavieja, G. G. (2014). idTracker: tracking individuals in a group by automatic identification of unmarked animals. *Nature methods*, 11(7):743–8.

Pievani, T. (2014). Individuals and groups in evolution: Darwinian pluralism and the multilevel selection debate. *Journal of Biosciences*, 39(April):319–325.

Plotkin, H. (1988). *The Role of Behavior in Evolution*. MIT Press.

Pollard, K. a. and Blumstein, D. T. (2011). Social group size predicts the evolution of individuality. *Current Biology*, 21(5):413–417.

Réale, D., Garant, D., Humphries, M. M., Bergeron, P., Careau, V., and Montiglio, P.-O. (2010). Personality and the emergence of the pace-of-life syndrome concept at the population level. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365:4051–4063.

Robinson, G. E., Fernald, R. D., and Clayton, D. F. (2008). Genes and social behavior. *Science*, 322(November):896–900.

Roeser, T. and Baier, H. (2003). Visuomotor behaviors in larval zebrafish after GFP-guided laser ablation of the optic tectum. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 23(9):3726–3734.

Romey, W. L. (1996). Individual differences make a difference in the trajectories of simulated schools of fish. *Ecological Modelling*, 92(1):65–77.

Saastamoinen, M., Ikonen, S., and Hanski, I. (2009). Significant effects of Pgi genotype and body reserves on lifespan in the Glanville fritillary butterfly. *Proceedings. Biological sciences / The Royal Society*, 276(January):1313–1322.

Santema, P. and Clutton-Brock, T. (2013). Meerkat helpers increase sentinel behaviour and bipedal vigilance in the presence of pups. *Animal Behaviour*, 85(3):655–661.

Schindler, D. E., Hilborn, R., Chasco, B., Boatright, C. P., Quinn, T. P., Rogers, L. a., and Webster, M. S. (2010). Population diversity and the portfolio effect in an exploited species. *Nature*, 465(7298):609–612.

Seeley, T. (1997). Honey bee colonies are group-level adaptive units. *The American Naturalist*, 150:S22–S41.

Seeley, T. and Buhrman, S. (1999). Group decision making in swarms of honey bees. *Behav. Ecol. Sociobiol.*, 45:19–31.

Seidel, G. J., Elsden, R., and Hasler, J. (2003). Embryo Transfer in the Dairy Herd. *Mississippi State University Extension Service*.

Seong, K. H., Li, D., Shimizu, H., Nakamura, R., and Ishii, S. (2011). Inheritance of stress-induced, ATF-2-dependent epigenetic change. *Cell*, 145(7):1049–1061.

Shapley, L. and Grofman, B. (1984). Optimizing group judgmental accuracy in the presence of interdependencies. *Public Choice*, 43:329–343.

Sharma, S. V., Lee, D. Y., Li, B., Quinlan, M. P., Takahashi, F., Maheswaran, S., Mcdermott, U., Azizian, N., Zou, L., Fischbach, M. a., Wong, K.-k., Brandstetter, K., Wittner, B., Ramaswamy, S., Classon, M., and Settleman, J. (2010). A chromatin-mediated reversible drug tolerant state in cancer cell subpopulations. *Cell*, 141(1):69–80.

Sheldon, M., Rice, D. S., D'Arcangelo, G., Yoneshima, H., Nakajima, K., Mikoshiba, K., Howell, B. W., Cooper, J. a., Goldowitz, D., and Curran, T. (1997). Scrambler and yotari disrupt the disabled gene and produce a reeler-like phenotype in mice. *Nature*, 389(October):730–733.

Shi, Y., Seto, E., Chang, L. S., and Shenk, T. (1991). Transcriptional repression by YY1, a human GLI-Krüppel-related protein, and relief of repression by adenovirus E1A protein. *Cell*, 67:377–388.

Shin, T., Kraemer, D., Pryor, J., Liu, L., Rugila, J., Howe, L., Buck, S., Murphy, K., Lyons, L., and Westhusin, M. (2002). A cat cloned by nuclear transplantation. *Nature*, 415:859.

Sibly, R. M. (1983). Optimal Group Size is unstable. *Animal Behaviour*, 31:947–948.

Smith, B. R. and Blumstein, D. T. (2008). Fitness consequences of personality: A meta-analysis. *Behavioral Ecology*, 19(January):448–455.

Stewart, K.J. and Harcourt, A. (1994). Gorillas vocalizations during rest periods.

Strahl, B. D. and Allis, C. D. (2000). The language of covalent histone modifications. *Nature*, 403:41–45.

Sumpter, D. J. T. (2010). *Collective Animal Behavior*.

Sumpter, D. J. T., Krause, J., James, R., Couzin, I. D., and Ward, A. J. W. (2008). Consensus decision making by fish. *Current biology : CB*, 18(22):1773–7.

Sumpter, D. J. T., Krause, J., Ward, A. J. W., Herbert-read, J. E., Sumpter, D. J. T., and Krause, J. (2011). Correction for Ward et al., Fast and accurate decisions through collective vigilance in fish shoals. *Proceedings of the National Academy of Sciences of the United States of America*, 108(6):E27.

Sumpter, D. J. T. and Pratt, S. C. (2009). Quorum responses and consensus decision making. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364(December 2008):743–753.

Surowiecki, J. (2004). *The Wisdom of Crowds*. Random House, Inc.

Takahashi, H., McCaffery, J. M., Irizarry, R. a., and Boeke, J. D. (2006). Nucleocytosolic Acetyl-Coenzyme A Synthetase Is Required for Histone Acetylation and Global Transcription. *Molecular Cell*, 23:207–217.

Thirumalai, V. and Cline, H. T. (2008). Endogenous dopamine suppresses initiation of swimming in prefeeding zebrafish larvae. *Journal of neurophysiology*, 100(June 2008):1635–1648.

Thorne, B. L. (1997). Evolution Of Eusociality In Termites. *Annu. Rev. Ecol. Syst*, (25):27–54.

Tibbetts, E. a. and Dale, J. (2007). Individual recognition: it is good to be different. *Trends in Ecology and Evolution*, 22(10):529–537.

Tinbergen, N. (1963). On aims and methods of Ethology. *Zeitschrift für Tierpsychologie*, 20(4):410–433.

Valor, L. M., Guiretti, D., Lopez-Atalaya, J. P., and Barco, A. (2013). Genomic landscape of transcriptional and epigenetic dysregulation in early onset polyglutamine disease. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 33(25):10471–82.

VanderWaal, K. L., Mosser, A., and Packer, C. (2009). Optimal group size, dispersal decisions and postdispersal relationships in female African lions. *Animal Behaviour*, 77(4):949–954.

Veitia, R. a. (2005). Stochasticity or the fatal 'imperfection' of cloning. *Journal of biosciences*, 30(1):21–30.

Viscido, S. V., Parrish, J. K., and Grünbaum, D. (2007). Factors influencing the structure and maintenance of fish schools. *Ecological Modelling*, 206(1-2):153–165.

Vogt, G. (2015). *Stochastic developmental variation, an epigenetic source of phenotypic diversity with far-reaching biological consequences*, volume 40.

Vogt, G., Huber, M., Thiemann, M., van den Boogaart, G., Schmitz, O. J., and Schubart, C. D. (2008). Production of different phenotypes from the same genotype in the same environment by developmental variation. *The Journal of experimental biology*, 211:510–523.

Vul, E. and Pashler, H. (2008). Measuring the Crowd Within: Probabilistic Representation Within Individuals. *Psych. Sci.*, 19(7):645–647.

Wagner, C. and Vinaimont, T. (2010). Evaluating the wisdom of crowds. *Proceedings of Issues in Information Systems*, XI(1):724–732.

Wang, D., Jao, L.-E., Zheng, N., Dolan, K., Ivey, J., Zonies, S., Wu, X., Wu, K., Yang, H., Meng, Q., Zhu, Z., Zhang, B., Lin, S., and Burgess, S. M. (2007). Efficient genome-wide mutagenesis of zebrafish genes by retroviral insertions.

Proceedings of the National Academy of Sciences of the United States of America, 104(30):12428–12433.

Wang, H., Duclot, F., Liu, Y., Wang, Z., and Kabbaj, M. (2013). Histone deacetylase inhibitors facilitate partner preference formation in female prairie voles. *Nature neuroscience*, 16(7):919–24.

Ward, A. J. W., Sumpter, D. J. T., Couzin, I. D., Hart, P. J. B., and Krause, J. (2008a). Quorum decision-making facilitates information transfer in fish shoals. *Proceedings of the National Academy of Sciences of the United States of America*, 105(19):6948–6953.

Ward, A. J. W., Sumpter, D. J. T., Couzin, I. D., Hart, P. J. B., and Krause, J. (2008b). Quorum decision-making facilitates information transfer in fish shoals. *Proceedings of the National Academy of Sciences of the United States of America*, 105(19):6948–53.

Ward, P. and Zahavi, a. (1973). The imporatance of certain assemblages of birds as "information-centres" for food-finding. *Ibis*, 115:517–534.

Wcislo, W. T. (1989). Behavioral Environments and Evolutionary Change. *Annual Review of Ecology and Systematics*, 20(1989):137–169.

Weigel, D. and Colot, V. (2012). Epialleles in plant evolution. *Genome Biology*, 13(10):249.

Weimerskirch, H., Martin, J., Clerquin, Y., Alexandre, P., and Jiraskova, S. (2001). Energy saving in flight formation. *Nature*, 413:697–698.

Weinstock, G. M., Robinson, G. E., and Gibbs, R. A. (2006). Insights into social insects from the genome of the honeybee Apis mellifera. *Nature*, 443(October):931–949.

Wernet, M. F., Mazzoni, E. O., Celik, A., Duncan, D. M., Duncan, I., and Desplan, C. (2006). Stochastic spineless expression creates the retinal mosaic for colour vision. *Nature*, 440(March):174–180.

White, T. C. R. (2008). The role of food, weather and climate in limiting the abundance of animals. *Biological Reviews*, 83:227–248.

Whitney, O., Pfenning, A. R., Howard, J. T., Blatti, C. a., Liu, F., Ward, J. M., Wang, R., Audet, J.-N., Kellis, M., Mukherjee, S., Sinha, S., Hartemink, A. J.,

West, A. E., and Jarvis, E. D. (2014). Core and region-enriched networks of behaviorally regulated genes and the singing genome. *Science*, 346(6215).

Williams, C. K., Lutz, R. S., and Applegate, R. D. (2003). Optimal group size and northern bobwhite coveys. *Animal Behaviour*, 66:377–387.

Wolf, M. and Weissing, F. J. (2012). Animal personalities: Consequences for ecology and evolution. *Trends in Ecology and Evolution*, 27(8):452–461.

Yao, Y.-l., Yang, W.-m., and Seto, E. (2001). Regulation of Transcription Factor YY1 by Acetylation and Deacetylation Regulation of Transcription Factor YY1 by Acetylation and Deacetylation. *Molecular and cellular biology*, 21(17):5979–5991.

Zöttl, M., Frommen, J. G., and Taborsky, M. (2013). Group size adjustment to ecological demand in a cooperative breeder. *Proceedings. Biological sciences / The Royal Society*, 280:20122772.

A. Genomic coordinates of hyper-variable acetylated regions

Genomic coordinates of hyper-variable acetylated regions in zebrafish larva. It also includes the nearest gene position for each region.

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
1	3709576	3709600	ENSDARG00000087467	3427787	3479522
1	3709601	3709625	ENSDARG00000087467	3427787	3479522
1	3709626	3709650	ENSDARG00000087467	3427787	3479522
1	12368776	12368800	ENSDARG00000089805	12145910	12155872
1	13428226	13428250	ENSDARG00000070424	13427656	13429261
1	13430201	13430225	ENSDARG00000090330	13428608	13428705
1	13441376	13441400	ENSDARG00000090249	13439641	13439738
1	13441476	13441500	ENSDARG00000090249	13439641	13439738
1	13442401	13442425	ENSDARG00000090249	13439641	13439738
1	13444851	13444875	ENSDARG00000090249	13439641	13439738
1	13444876	13444900	ENSDARG00000090249	13439641	13439738
1	13444901	13444925	ENSDARG00000090249	13439641	13439738
1	13458901	13458925	ENSDARG00000091356	13460601	13460698
1	13461426	13461450	ENSDARG00000088897	13482169	13482266
1	13485501	13485525	ENSDARG00000088602	13481619	13483218
1	13485526	13485550	ENSDARG00000088602	13481619	13483218
1	16713976	16714000	ENSDARG00000003046	16489551	16545516
1	16714001	16714025	ENSDARG00000003046	16489551	16545516
1	16714401	16714425	ENSDARG00000003046	16489551	16545516
1	16714426	16714450	ENSDARG00000003046	16489551	16545516
1	16912251	16912275	ENSDARG00000090504	16908287	16908405
1	16912301	16912325	ENSDARG00000090504	16908287	16908405

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
1	16912326	16912350	ENSDARG00000090504	16908287	16908405
1	20533176	20533200	ENSDARG00000055443	20393123	20548992
1	23457851	23457875	ENSDARG00000037309	23377312	23415647
1	23457876	23457900	ENSDARG00000037309	23377312	23415647
1	23458026	23458050	ENSDARG00000037309	23377312	23415647
1	23458051	23458075	ENSDARG00000037309	23377312	23415647
1	23458076	23458100	ENSDARG00000037309	23377312	23415647
1	26838576	26838600	ENSDARG00000069996	26837431	26863161
1	26838601	26838625	ENSDARG00000069996	26837431	26863161
1	37172601	37172625	ENSDARG00000036595	37162968	37212687
1	37172626	37172650	ENSDARG00000036595	37162968	37212687
1	37172651	37172675	ENSDARG00000036595	37162968	37212687
1	37172676	37172700	ENSDARG00000036595	37162968	37212687
1	44118376	44118400	ENSDARG00000052696	44096669	44110096
1	44118401	44118425	ENSDARG00000052696	44096669	44110096
1	44702926	44702950	ENSDARG00000059682	44693717	44708951
1	44702951	44702975	ENSDARG00000059682	44693717	44708951
1	44702976	44703000	ENSDARG00000059682	44693717	44708951
1	44703001	44703025	ENSDARG00000059682	44693717	44708951
1	49454626	49454650	ENSDARG00000096191	49428343	49428672
2	3175451	3175475	ENSDARG00000086151	3385910	3386762
2	3175476	3175500	ENSDARG00000086151	3385910	3386762
2	3891226	3891250	ENSDARG00000063445	3879432	3898088
2	3891251	3891275	ENSDARG00000063445	3879432	3898088
2	5542976	5543000	ENSDARG00000056478	5530274	5565810
2	7559701	7559725	ENSDARG00000008966	7552562	7585043
2	7559726	7559750	ENSDARG00000008966	7552562	7585043
2	8278501	8278525	ENSDARG00000091887	8274929	8279057
2	13141701	13141725	ENSDARG00000095837	13090040	13125921
2	14078851	14078875	ENSDARG00000074233	13915648	14057782
2	14078876	14078900	ENSDARG00000074233	13915648	14057782
2	14078901	14078925	ENSDARG00000074233	13915648	14057782
2	14078926	14078950	ENSDARG00000074233	13915648	14057782
2	14485551	14485575	ENSDARG00000080143	14485751	14485936
2	14485601	14485625	ENSDARG00000080143	14485751	14485936
2	14501051	14501075	ENSDARG00000083104	14501191	14501376

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
2	21704351	21704375	ENSDARG00000020261	21618206	21644458
2	21704376	21704400	ENSDARG00000020261	21618206	21644458
2	26801801	26801825	ENSDARG00000055945	26804453	26830352
2	26801826	26801850	ENSDARG00000055945	26804453	26830352
2	26801851	26801875	ENSDARG00000055945	26804453	26830352
2	31839851	31839875	ENSDARG00000094081	31833726	31837071
2	32322651	32322675	ENSDARG00000094100	32322722	32325902
2	32322676	32322700	ENSDARG00000094100	32322722	32325902
2	38121526	38121550	ENSDARG00000079605	38045045	38065775
2	38121551	38121575	ENSDARG00000079605	38045045	38065775
2	45184776	45184800	ENSDARG00000038585	45178320	45197119
2	58512076	58512100	ENSDARG00000094356	58439050	58448431
2	58512101	58512125	ENSDARG00000094356	58439050	58448431
3	7546251	7546275	ENSDARG00000087816	7520664	7565045
3	7792851	7792875	ENSDARG00000091679	7791025	7806093
3	9542376	9542400	ENSDARG00000091688	9544448	9564878
3	21175776	21175800	ENSDARG00000006566	21156841	21184991
3	21175801	21175825	ENSDARG00000006566	21156841	21184991
3	23635101	23635125	ENSDARG00000079912	23621224	23698205
3	23635126	23635150	ENSDARG00000079912	23621224	23698205
3	23635151	23635175	ENSDARG00000079912	23621224	23698205
3	30256026	30256050	ENSDARG00000077053	30193895	30251434
3	30256051	30256075	ENSDARG00000077053	30193895	30251434
3	30256076	30256100	ENSDARG00000077053	30193895	30251434
3	31626151	31626175	ENSDARG00000054446	31737513	31742339
3	31626226	31626250	ENSDARG00000054446	31737513	31742339
3	33067751	33067775	ENSDARG00000014646	33061160	33071335
3	33067776	33067800	ENSDARG00000014646	33061160	33071335
3	33067801	33067825	ENSDARG00000014646	33061160	33071335
3	34700526	34700550	ENSDARG00000076437	34687148	34698397
3	34700551	34700575	ENSDARG00000076437	34687148	34698397
3	37617026	37617050	ENSDARG00000053070	37644218	37655642
3	38967026	38967050	ENSDARG00000095329	38966483	38970137
3	44704976	44705000	ENSDARG00000052648	44579435	44791438
3	44705001	44705025	ENSDARG00000052648	44579435	44791438
3	46577001	46577025	ENSDARG00000020326	46518026	46590220

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
3	46577026	46577050	ENSDARG00000020326	46518026	46590220
3	46577051	46577075	ENSDARG00000020326	46518026	46590220
3	49437626	49437650	ENSDARG00000040099	49394281	49489234
3	49437651	49437675	ENSDARG00000040099	49394281	49489234
3	49437676	49437700	ENSDARG00000040099	49394281	49489234
3	54974276	54974300	ENSDARG00000077847	55019428	55088977
3	58117501	58117525	ENSDARG00000087367	58120130	58121254
3	58117551	58117575	ENSDARG00000087367	58120130	58121254
3	58117576	58117600	ENSDARG00000087367	58120130	58121254
3	58120876	58120900	ENSDARG00000087367	58120130	58121254
3	59411426	59411450	ENSDARG00000031618	59375244	59393109
3	61548951	61548975	ENSDARG00000074669	61554065	61573732
3	61548976	61549000	ENSDARG00000074669	61554065	61573732
3	61549001	61549025	ENSDARG00000074669	61554065	61573732
3	61549026	61549050	ENSDARG00000074669	61554065	61573732
4	457751	457775	ENSDARG00000045900	444235	467256
4	1630651	1630675	ENSDARG00000070477	1630121	1645108
4	1630676	1630700	ENSDARG00000070477	1630121	1645108
4	1630701	1630725	ENSDARG00000070477	1630121	1645108
4	1630726	1630750	ENSDARG00000070477	1630121	1645108
4	2133876	2133900	ENSDARG00000053709	2128621	2142410
4	2133901	2133925	ENSDARG00000053709	2128621	2142410
4	7448776	7448800	ENSDARG00000093557	7448258	7449637
4	10572826	10572850	ENSDARG00000045765	10547885	10589412
4	10572851	10572875	ENSDARG00000045765	10547885	10589412
4	10572876	10572900	ENSDARG00000045765	10547885	10589412
4	10572901	10572925	ENSDARG00000045765	10547885	10589412
4	14671501	14671525	ENSDARG00000004336	14671236	14881932
4	14671526	14671550	ENSDARG00000004336	14671236	14881932
4	14671551	14671575	ENSDARG00000004336	14671236	14881932
4	15562126	15562150	ENSDARG00000088002	15514306	15518163
4	19442051	19442075	ENSDARG00000045485	19440535	19457131
4	19442076	19442100	ENSDARG00000045485	19440535	19457131
4	22028726	22028750	ENSDARG00000021590	21694608	22018675
4	22028751	22028775	ENSDARG00000021590	21694608	22018675
4	28270851	28270875	ENSDARG00000088214	28270325	28284111

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
4	28270876	28270900	ENSDARG00000088214	28270325	28284111
4	28680851	28680875	ENSDARG00000076054	28611145	28844703
4	29824326	29824350	ENSDARG00000096117	29817206	29822914
4	29969776	29969800	ENSDARG00000096192	29987295	29999211
4	30340601	30340625	ENSDARG00000087338	30347638	30491764
4	30767401	30767425	ENSDARG00000084499	30749428	30749529
4	30776476	30776500	ENSDARG00000090539	30771549	30776372
4	31219626	31219650	ENSDARG00000096005	31175908	31176505
4	31663051	31663075	ENSDARG00000095997	31590932	31606407
4	31663076	31663100	ENSDARG00000095997	31590932	31606407
4	31663101	31663125	ENSDARG00000095997	31590932	31606407
4	32007076	32007100	ENSDARG00000090220	31941329	32042427
4	32740201	32740225	ENSDARG00000087544	32698101	32699827
4	32740226	32740250	ENSDARG00000087544	32698101	32699827
4	33796301	33796325	ENSDARG00000092293	33824429	33829364
4	36238951	36238975	ENSDARG00000092157	36272503	36273366
4	36238976	36239000	ENSDARG00000092157	36272503	36273366
4	36590826	36590850	ENSDARG00000095668	36569895	36578464
4	36590851	36590875	ENSDARG00000095668	36569895	36578464
4	36706476	36706500	ENSDARG00000086791	36646547	36846863
4	36706501	36706525	ENSDARG00000086791	36646547	36846863
4	39017501	39017525	ENSDARG00000094269	39032959	39038703
4	39017526	39017550	ENSDARG00000094269	39032959	39038703
4	39017551	39017575	ENSDARG00000094269	39032959	39038703
4	39387251	39387275	ENSDARG00000094017	39381777	39398900
4	39387276	39387300	ENSDARG00000094017	39381777	39398900
4	39387301	39387325	ENSDARG00000094017	39381777	39398900
4	39387326	39387350	ENSDARG00000094017	39381777	39398900
4	39456726	39456750	ENSDARG00000080979	39451275	39451423
4	39456751	39456775	ENSDARG00000080979	39451275	39451423
4	39947126	39947150	ENSDARG00000096166	39939590	39947909
4	39947151	39947175	ENSDARG00000096166	39939590	39947909
4	42309651	42309675	ENSDARG00000091811	42217858	42268068
4	43374451	43374475	ENSDARG00000096476	43368311	43375347
4	43374476	43374500	ENSDARG00000096476	43368311	43375347
4	43374501	43374525	ENSDARG00000096476	43368311	43375347

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
4	43374526	43374550	ENSDARG00000096476	43368311	43375347
4	43386926	43386950	ENSDARG00000096485	43375089	43380828
4	43462801	43462825	ENSDARG00000087222	43450550	43465716
4	43462826	43462850	ENSDARG00000087222	43450550	43465716
4	43462851	43462875	ENSDARG00000087222	43450550	43465716
4	43898576	43898600	ENSDARG00000090331	43892184	43892281
4	45536201	45536225	ENSDARG00000087626	45514256	45518651
4	45892026	45892050	ENSDARG00000089587	45853746	45939686
4	45892051	45892075	ENSDARG00000089587	45853746	45939686
4	45970601	45970625	ENSDARG00000089587	45853746	45939686
4	45970626	45970650	ENSDARG00000089587	45853746	45939686
4	45970651	45970675	ENSDARG00000089587	45853746	45939686
4	45970676	45970700	ENSDARG00000089587	45853746	45939686
4	46506876	46506900	ENSDARG00000095194	46508264	46509771
4	48239901	48239925	ENSDARG00000085993	48240956	48241062
4	48716276	48716300	ENSDARG00000077485	48718445	48786306
4	48716301	48716325	ENSDARG00000077485	48718445	48786306
4	51642651	51642675	ENSDARG00000092794	51617439	51655969
4	51642676	51642700	ENSDARG00000092794	51617439	51655969
4	53390176	53390200	ENSDARG00000096065	53431494	53432677
4	53390201	53390225	ENSDARG00000096065	53431494	53432677
4	53461326	53461350	ENSDARG00000094888	53456594	53464880
4	53461351	53461375	ENSDARG00000094888	53456594	53464880
4	53461376	53461400	ENSDARG00000094888	53456594	53464880
4	53620626	53620650	ENSDARG00000023991	53659156	53660585
4	54747576	54747600	ENSDARG00000089363	54714148	54781135
4	54747601	54747625	ENSDARG00000089363	54714148	54781135
4	56009701	56009725	ENSDARG00000096024	55986432	55990645
4	56009726	56009750	ENSDARG00000096024	55986432	55990645
4	56230226	56230250	ENSDARG00000088301	56230176	56260016
4	56230251	56230275	ENSDARG00000088301	56230176	56260016
4	56998776	56998800	ENSDARG00000096211	56973619	56977280
4	56998801	56998825	ENSDARG00000096211	56973619	56977280
4	61956326	61956350	ENSDARG00000088303	61954567	61960605
4	61956351	61956375	ENSDARG00000088303	61954567	61960605
4	61956376	61956400	ENSDARG00000088303	61954567	61960605

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
4	61956401	61956425	ENSDARG00000088303	61954567	61960605
4	61956426	61956450	ENSDARG00000088303	61954567	61960605
5	1184876	1184900	ENSDARG00000086896	1219724	1227129
5	1184901	1184925	ENSDARG00000086896	1219724	1227129
5	1185751	1185775	ENSDARG00000086896	1219724	1227129
5	14213651	14213675	ENSDARG00000002300	14011701	14026222
5	14213676	14213700	ENSDARG00000002300	14011701	14026222
5	14213701	14213725	ENSDARG00000002300	14011701	14026222
5	14213726	14213750	ENSDARG00000002300	14011701	14026222
5	18132301	18132325	ENSDARG00000090174	18067321	18083466
5	18132326	18132350	ENSDARG00000090174	18067321	18083466
5	18132351	18132375	ENSDARG00000090174	18067321	18083466
5	29498126	29498150	ENSDARG00000035514	29542775	29565255
5	29498151	29498175	ENSDARG00000035514	29542775	29565255
5	29498176	29498200	ENSDARG00000035514	29542775	29565255
5	34242901	34242925	ENSDARG00000090214	34255546	34255663
5	34242926	34242950	ENSDARG00000090214	34255546	34255663
5	34242951	34242975	ENSDARG00000090214	34255546	34255663
5	35798751	35798775	ENSDARG00000060393	35684489	35803395
5	35798776	35798800	ENSDARG00000060393	35684489	35803395
5	35798801	35798825	ENSDARG00000060393	35684489	35803395
5	43997626	43997650	ENSDARG00000067762	43998099	44001830
5	44002101	44002125	ENSDARG00000094728	43999855	44009022
5	44002126	44002150	ENSDARG00000094728	43999855	44009022
5	44002301	44002325	ENSDARG00000094728	43999855	44009022
5	44009451	44009475	ENSDARG00000093543	44007046	44016213
5	44023501	44023525	ENSDARG00000095645	44021118	44030279
5	44023526	44023550	ENSDARG00000095645	44021118	44030279
5	44025951	44025975	ENSDARG00000095645	44021118	44030279
5	44026826	44026850	ENSDARG00000095645	44021118	44030279
5	44032851	44032875	ENSDARG00000091034	44027940	44028037
5	44032876	44032900	ENSDARG00000091034	44027940	44028037
5	44032901	44032925	ENSDARG00000091034	44027940	44028037
5	44033151	44033175	ENSDARG00000091034	44027940	44028037
5	44092026	44092050	ENSDARG00000090463	44086823	44086920
5	44097576	44097600	ENSDARG00000089623	44094018	44094115

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
5	44099526	44099550	ENSDARG00000089623	44094018	44094115
5	44103851	44103875	ENSDARG00000091188	44101221	44101318
5	44103976	44104000	ENSDARG00000091188	44101221	44101318
5	44104001	44104025	ENSDARG00000091188	44101221	44101318
5	44104026	44104050	ENSDARG00000091188	44101221	44101318
5	44104126	44104150	ENSDARG00000091188	44101221	44101318
5	44106426	44106450	ENSDARG00000091188	44101221	44101318
5	44111176	44111200	ENSDARG00000090917	44108417	44108514
5	44114601	44114625	ENSDARG00000090917	44108417	44108514
5	44118376	44118400	ENSDARG00000087541	44115624	44115721
5	44118401	44118425	ENSDARG00000087541	44115624	44115721
5	44120551	44120575	ENSDARG00000087541	44115624	44115721
5	44120576	44120600	ENSDARG00000087541	44115624	44115721
5	44120601	44120625	ENSDARG00000087541	44115624	44115721
5	44120626	44120650	ENSDARG00000087541	44115624	44115721
5	44121751	44121775	ENSDARG00000087541	44115624	44115721
5	44121851	44121875	ENSDARG00000087541	44115624	44115721
5	44121876	44121900	ENSDARG00000087541	44115624	44115721
5	44125451	44125475	ENSDARG00000091336	44122821	44122918
5	44125576	44125600	ENSDARG00000091336	44122821	44122918
5	44125626	44125650	ENSDARG00000091336	44122821	44122918
5	44126251	44126275	ENSDARG00000091336	44122821	44122918
5	44274101	44274125	ENSDARG00000088892	44278358	44278455
5	44276826	44276850	ENSDARG00000067767	44276230	44278956
5	44287576	44287600	ENSDARG00000089460	44292738	44292835
5	44287601	44287625	ENSDARG00000089460	44292738	44292835
5	44289001	44289025	ENSDARG00000089460	44292738	44292835
5	44289426	44289450	ENSDARG00000089460	44292738	44292835
5	44290001	44290025	ENSDARG00000089460	44292738	44292835
5	44290026	44290050	ENSDARG00000089460	44292738	44292835
5	45767401	45767425	ENSDARG00000094625	45683346	45809897
5	45767426	45767450	ENSDARG00000094625	45683346	45809897
5	46739476	46739500	ENSDARG00000008904	46757516	46902718
5	47198776	47198800	ENSDARG00000059963	47506898	47525564
5	47198801	47198825	ENSDARG00000059963	47506898	47525564
5	48907026	48907050	ENSDARG00000093413	48273390	48507648

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
5	51349976	51350000	ENSDARG00000052695	51527504	51534340
5	54528726	54528750	ENSDARG00000086999	54412566	54549720
5	57349526	57349550	ENSDARG00000035184	57343313	57361312
5	58588176	58588200	ENSDARG00000059669	58415371	58611483
5	58588226	58588250	ENSDARG00000059669	58415371	58611483
5	59887426	59887450	ENSDARG00000067517	59888458	59917195
5	59887451	59887475	ENSDARG00000067517	59888458	59917195
5	66494151	66494175	ENSDARG00000094107	66488818	66494803
5	66494176	66494200	ENSDARG00000094107	66488818	66494803
5	74916026	74916050	ENSDARG00000003803	74915009	74932802
5	74916051	74916075	ENSDARG00000003803	74915009	74932802
5	74916076	74916100	ENSDARG00000003803	74915009	74932802
5	74916101	74916125	ENSDARG00000003803	74915009	74932802
6	6288751	6288775	ENSDARG00000078440	6218335	6262806
6	6288776	6288800	ENSDARG00000078440	6218335	6262806
6	6300526	6300550	ENSDARG00000078440	6218335	6262806
6	6300551	6300575	ENSDARG00000078440	6218335	6262806
6	10104576	10104600	ENSDARG00000071113	10080895	10144304
6	13137276	13137300	ENSDARG00000028546	13116491	13141926
6	15805276	15805300	ENSDARG00000057276	15662473	15810851
6	15805301	15805325	ENSDARG00000057276	15662473	15810851
6	15805326	15805350	ENSDARG00000057276	15662473	15810851
6	19684951	19684975	ENSDARG00000052589	19682034	19706928
6	19684976	19685000	ENSDARG00000052589	19682034	19706928
6	24639201	24639225	ENSDARG00000091125	24643486	24643583
6	24678251	24678275	ENSDARG00000093631	24642917	24645498
6	24678276	24678300	ENSDARG00000093631	24642917	24645498
6	24719601	24719625	ENSDARG00000090260	24749737	24752211
6	24719626	24719650	ENSDARG00000090260	24749737	24752211
6	24719651	24719675	ENSDARG00000090260	24749737	24752211
6	25005576	25005600	ENSDARG00000088032	25025336	25025433
6	28886526	28886550	ENSDARG00000017886	28858572	28873684
6	35350801	35350825	ENSDARG00000070074	35013405	35311477
6	35400326	35400350	ENSDARG00000070074	35013405	35311477
6	35511001	35511025	ENSDARG00000070013	35599989	35606659
6	35520351	35520375	ENSDARG00000070013	35599989	35606659

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
6	36061226	36061250	ENSDARG00000007080	36160310	36170249
6	36064951	36064975	ENSDARG00000007080	36160310	36170249
6	36064976	36065000	ENSDARG00000007080	36160310	36170249
6	36065001	36065025	ENSDARG00000007080	36160310	36170249
6	41569151	41569175	ENSDARG00000060309	41486570	41689273
6	44107876	44107900	ENSDARG00000044191	44095091	44125036
6	44792176	44792200	ENSDARG00000076233	44850988	44916638
6	44792201	44792225	ENSDARG00000076233	44850988	44916638
6	50546176	50546200	ENSDARG00000059604	50532243	50637165
6	52988001	52988025	ENSDARG00000061222	52900365	52991541
6	52988026	52988050	ENSDARG00000061222	52900365	52991541
6	52988051	52988075	ENSDARG00000061222	52900365	52991541
6	53833401	53833425	ENSDARG00000039182	53453187	53841950
6	53833426	53833450	ENSDARG00000039182	53453187	53841950
6	53833451	53833475	ENSDARG00000039182	53453187	53841950
6	53833476	53833500	ENSDARG00000039182	53453187	53841950
7	5496501	5496525	ENSDARG00000013743	5463775	5492137
7	7612101	7612125	ENSDARG00000045199	7489102	7559198
7	7612151	7612175	ENSDARG00000045199	7489102	7559198
7	8680951	8680975	ENSDARG00000060176	8604057	8616573
7	17284626	17284650	ENSDARG00000025868	17280340	17322115
7	17284651	17284675	ENSDARG00000025868	17280340	17322115
7	17284676	17284700	ENSDARG00000025868	17280340	17322115
7	30378376	30378400	ENSDARG00000036055	30377981	30379395
7	34505876	34505900	ENSDARG00000095082	34506795	34509817
7	35502251	35502275	ENSDARG00000036096	35450190	35493955
7	35782101	35782125	ENSDARG00000043170	35756118	35851568
7	35921826	35921850	ENSDARG00000077584	35942441	35947151
7	40384726	40384750	ENSDARG00000014439	40354309	40545807
7	42445376	42445400	ENSDARG00000052167	42231105	42254717
7	42452226	42452250	ENSDARG00000052167	42231105	42254717
7	42582976	42583000	ENSDARG00000088005	42708713	42708827
7	42583001	42583025	ENSDARG00000088005	42708713	42708827
7	46474651	46474675	ENSDARG00000077402	46580356	46583100
7	46474676	46474700	ENSDARG00000077402	46580356	46583100
7	47203701	47203725	ENSDARG00000086413	47322222	47322336

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
7	47203726	47203750	ENSDARG00000086413	47322222	47322336
7	47203751	47203775	ENSDARG00000086413	47322222	47322336
7	47203776	47203800	ENSDARG00000086413	47322222	47322336
7	48262501	48262525	ENSDARG00000091033	48517958	48518072
7	48262526	48262550	ENSDARG00000091033	48517958	48518072
7	48262976	48263000	ENSDARG00000091033	48517958	48518072
7	50198601	50198625	ENSDARG00000028058	50179826	50195706
7	50198626	50198650	ENSDARG00000028058	50179826	50195706
7	50198651	50198675	ENSDARG00000028058	50179826	50195706
7	50198676	50198700	ENSDARG00000028058	50179826	50195706
7	56502326	56502350	ENSDARG00000089538	56487218	56555379
7	57683101	57683125	ENSDARG00000078882	57690264	57732027
7	58036326	58036350	ENSDARG00000073856	58033636	58035408
7	58036351	58036375	ENSDARG00000073856	58033636	58035408
7	59893401	59893425	ENSDARG00000036045	59931811	59941037
7	63132251	63132275	ENSDARG00000091417	63113745	63113859
7	65695126	65695150	ENSDARG00000068210	65672243	65681470
7	69359651	69359675	ENSDARG00000089723	69434673	69434789
7	69359676	69359700	ENSDARG00000089723	69434673	69434789
7	70554376	70554400	ENSDARG00000073681	70493881	70652404
7	70554401	70554425	ENSDARG00000073681	70493881	70652404
7	71955576	71955600	ENSDARG00000092441	71942524	71944217
7	71955601	71955625	ENSDARG00000092441	71942524	71944217
7	71955626	71955650	ENSDARG00000092441	71942524	71944217
7	72760501	72760525	ENSDARG00000090305	72827097	72827786
7	72760626	72760650	ENSDARG00000090305	72827097	72827786
7	72760651	72760675	ENSDARG00000090305	72827097	72827786
7	72760676	72760700	ENSDARG00000090305	72827097	72827786
7	72760701	72760725	ENSDARG00000090305	72827097	72827786
7	75511476	75511500	ENSDARG00000041086	75501697	75523464
7	75511501	75511525	ENSDARG00000041086	75501697	75523464
8	2865176	2865200	ENSDARG00000025846	2831299	2918891
8	2865201	2865225	ENSDARG00000025846	2831299	2918891
8	8309826	8309850	ENSDARG00000092239	8280204	8300077
8	13263176	13263200	ENSDARG00000093851	13038884	13186623
8	14274926	14274950	ENSDARG00000078567	14217556	14304260

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
8	15577401	15577425	ENSDARG00000020131	15503957	15597524
8	19584051	19584075	ENSDARG00000061976	19456460	19576851
8	23919701	23919725	ENSDARG00000092039	23893670	23934610
8	27666401	27666425	ENSDARG00000092035	27671389	27691788
8	27666426	27666450	ENSDARG00000092035	27671389	27691788
8	27666451	27666475	ENSDARG00000092035	27671389	27691788
8	28030801	28030825	ENSDARG00000095503	28029819	28033808
8	28742276	28742300	ENSDARG00000056091	28675955	28711408
8	31686276	31686300	ENSDARG00000042428	31635153	31639308
8	31689901	31689925	ENSDARG00000042428	31635153	31639308
8	31690276	31690300	ENSDARG00000042428	31635153	31639308
8	31690301	31690325	ENSDARG00000042428	31635153	31639308
8	31707951	31707975	ENSDARG00000033871	31769445	31777680
8	31707976	31708000	ENSDARG00000033871	31769445	31777680
8	35267451	35267475	ENSDARG00000094032	35268611	35269813
8	35267476	35267500	ENSDARG00000094032	35268611	35269813
8	35268376	35268400	ENSDARG00000094032	35268611	35269813
8	35268401	35268425	ENSDARG00000094032	35268611	35269813
8	35271326	35271350	ENSDARG00000094717	35273318	35274520
8	35271351	35271375	ENSDARG00000094717	35273318	35274520
8	35271376	35271400	ENSDARG00000094717	35273318	35274520
8	35271401	35271425	ENSDARG00000094717	35273318	35274520
8	35276051	35276075	ENSDARG00000093481	35278029	35279541
8	35276226	35276250	ENSDARG00000093481	35278029	35279541
8	35276276	35276300	ENSDARG00000093481	35278029	35279541
8	35276301	35276325	ENSDARG00000093481	35278029	35279541
8	35277076	35277100	ENSDARG00000093481	35278029	35279541
8	35277101	35277125	ENSDARG00000093481	35278029	35279541
8	35277126	35277150	ENSDARG00000093481	35278029	35279541
8	35280901	35280925	ENSDARG00000090183	35171216	35256670
8	35305251	35305275	ENSDARG00000092675	35305988	35312699
8	35305276	35305300	ENSDARG00000092675	35305988	35312699
8	35313201	35313225	ENSDARG00000091883	35316215	35317727
8	35313226	35313250	ENSDARG00000091883	35316215	35317727
8	35314601	35314625	ENSDARG00000091883	35316215	35317727
8	35354901	35354925	ENSDARG00000095028	35356798	35358000

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
8	35355326	35355350	ENSDARG00000095028	35356798	35358000
8	35355351	35355375	ENSDARG00000095028	35356798	35358000
8	35359551	35359575	ENSDARG00000094765	35361513	35362715
8	35360601	35360625	ENSDARG00000094765	35361513	35362715
8	35360626	35360650	ENSDARG00000094765	35361513	35362715
8	35361776	35361800	ENSDARG00000094765	35361513	35362715
8	35364151	35364175	ENSDARG00000093508	35392542	35393744
8	35364176	35364200	ENSDARG00000093508	35392542	35393744
8	35365126	35365150	ENSDARG00000093508	35392542	35393744
8	35365176	35365200	ENSDARG00000093508	35392542	35393744
8	35367551	35367575	ENSDARG00000093508	35392542	35393744
8	35367576	35367600	ENSDARG00000093508	35392542	35393744
8	35367751	35367775	ENSDARG00000093508	35392542	35393744
8	35368851	35368875	ENSDARG00000093508	35392542	35393744
8	35368876	35368900	ENSDARG00000093508	35392542	35393744
8	35368901	35368925	ENSDARG00000093508	35392542	35393744
8	35369401	35369425	ENSDARG00000093508	35392542	35393744
8	35369426	35369450	ENSDARG00000093508	35392542	35393744
8	35387176	35387200	ENSDARG00000093508	35392542	35393744
8	35387201	35387225	ENSDARG00000093508	35392542	35393744
8	35387226	35387250	ENSDARG00000093508	35392542	35393744
8	35387376	35387400	ENSDARG00000093508	35392542	35393744
8	35387426	35387450	ENSDARG00000093508	35392542	35393744
8	35387551	35387575	ENSDARG00000093508	35392542	35393744
8	35389201	35389225	ENSDARG00000093508	35392542	35393744
8	35389226	35389250	ENSDARG00000093508	35392542	35393744
8	35391876	35391900	ENSDARG00000093508	35392542	35393744
8	35392026	35392050	ENSDARG00000093508	35392542	35393744
8	35395626	35395650	ENSDARG00000092670	35397222	35398417
8	35396676	35396700	ENSDARG00000092670	35397222	35398417
8	35396701	35396725	ENSDARG00000092670	35397222	35398417
8	35397551	35397575	ENSDARG00000092670	35397222	35398417
8	35397576	35397600	ENSDARG00000092670	35397222	35398417
8	35466376	35466400	ENSDARG00000093654	35468254	35469456
8	35469976	35470000	ENSDARG00000094733	35472962	35474164
8	35470101	35470125	ENSDARG00000094733	35472962	35474164

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
8	35470126	35470150	ENSDARG00000094733	35472962	35474164
8	35472051	35472075	ENSDARG00000094733	35472962	35474164
8	35472226	35472250	ENSDARG00000094733	35472962	35474164
8	35474651	35474675	ENSDARG00000094864	35477637	35479142
8	35474676	35474700	ENSDARG00000094864	35477637	35479142
8	35583126	35583150	ENSDARG00000093439	35586001	35587203
8	35583151	35583175	ENSDARG00000093439	35586001	35587203
8	35583176	35583200	ENSDARG00000093439	35586001	35587203
8	35583201	35583225	ENSDARG00000093439	35586001	35587203
8	35585076	35585100	ENSDARG00000093439	35586001	35587203
8	35585276	35585300	ENSDARG00000093439	35586001	35587203
8	35585301	35585325	ENSDARG00000093439	35586001	35587203
8	35738676	35738700	ENSDARG00000094167	35744756	35745958
8	35738701	35738725	ENSDARG00000094167	35744756	35745958
8	35752226	35752250	ENSDARG00000095172	35754186	35755388
8	35753251	35753275	ENSDARG00000095172	35754186	35755388
8	35753276	35753300	ENSDARG00000095172	35754186	35755388
8	35802751	35802775	ENSDARG00000094532	35810242	35811444
8	35805701	35805725	ENSDARG00000094532	35810242	35811444
8	35810801	35810825	ENSDARG00000094532	35810242	35811444
8	35819126	35819150	ENSDARG00000056266	35820268	35820911
8	35853651	35853675	ENSDARG00000090522	35854375	35855887
8	35854701	35854725	ENSDARG00000090522	35854375	35855887
8	39976201	39976225	ENSDARG00000060518	39738244	39846811
8	39976226	39976250	ENSDARG00000060518	39738244	39846811
8	40008726	40008750	ENSDARG00000060518	39738244	39846811
8	46190226	46190250	ENSDARG00000001154	46121698	46195595
8	46190251	46190275	ENSDARG00000001154	46121698	46195595
8	46190276	46190300	ENSDARG00000001154	46121698	46195595
8	54975226	54975250	ENSDARG00000035423	54975726	54978405
8	54975251	54975275	ENSDARG00000035423	54975726	54978405
8	54975276	54975300	ENSDARG00000035423	54975726	54978405
9	3537201	3537225	ENSDARG00000042467	3544979	3605958
9	3537226	3537250	ENSDARG00000042467	3544979	3605958
9	4345226	4345250	ENSDARG00000078117	4235364	4402553
9	4345251	4345275	ENSDARG00000078117	4235364	4402553

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
9	4345276	4345300	ENSDARG00000078117	4235364	4402553
9	4345301	4345325	ENSDARG00000078117	4235364	4402553
9	14076326	14076350	ENSDARG00000058207	14009788	14076449
9	14076351	14076375	ENSDARG00000058207	14009788	14076449
9	14076376	14076400	ENSDARG00000058207	14009788	14076449
9	16818351	16818375	ENSDARG00000039373	16828231	16919914
9	20187076	20187100	ENSDARG00000069936	20210249	20228109
9	20187101	20187125	ENSDARG00000069936	20210249	20228109
9	20194651	20194675	ENSDARG00000069936	20210249	20228109
9	20800426	20800450	ENSDARG00000093349	20714914	20796146
9	20800451	20800475	ENSDARG00000093349	20714914	20796146
9	20833501	20833525	ENSDARG00000093419	20854802	20856959
9	24462876	24462900	ENSDARG00000095527	24416759	24421420
9	28085176	28085200	ENSDARG00000093542	28334222	28335123
9	28085201	28085225	ENSDARG00000093542	28334222	28335123
9	30783451	30783475	ENSDARG00000062045	30749688	31018698
9	30783476	30783500	ENSDARG00000062045	30749688	31018698
9	30783501	30783525	ENSDARG00000062045	30749688	31018698
9	32001876	32001900	ENSDARG00000069440	31853598	32096012
9	32290801	32290825	ENSDARG00000029729	32201019	32292453
9	32290826	32290850	ENSDARG00000029729	32201019	32292453
9	33461051	33461075	ENSDARG00000067676	33391822	33494157
9	33461076	33461100	ENSDARG00000067676	33391822	33494157
9	33461101	33461125	ENSDARG00000067676	33391822	33494157
9	33461126	33461150	ENSDARG00000067676	33391822	33494157
9	34558426	34558450	ENSDARG00000059809	34399991	34510918
9	39563301	39563325	ENSDARG00000061265	39554995	39575190
9	39563326	39563350	ENSDARG00000061265	39554995	39575190
9	39630951	39630975	ENSDARG00000055136	39613770	39639338
9	44586676	44586700	ENSDARG00000006065	44313629	44611371
9	44636001	44636025	ENSDARG00000014008	44625558	44717304
9	44636026	44636050	ENSDARG00000014008	44625558	44717304
9	52779251	52779275	ENSDARG00000004712	52776692	52782926
9	52779276	52779300	ENSDARG00000004712	52776692	52782926
9	55165026	55165050	ENSDARG00000006008	55311304	55336342
9	55165051	55165075	ENSDARG00000006008	55311304	55336342

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
10	541926	541950	ENSDARG00000045632	565880	569157
10	7837701	7837725	ENSDARG00000053869	7814066	7837635
10	7837726	7837750	ENSDARG00000053869	7814066	7837635
10	11887701	11887725	ENSDARG00000069420	11756201	11806249
10	11887726	11887750	ENSDARG00000069420	11756201	11806249
10	11887751	11887775	ENSDARG00000069420	11756201	11806249
10	11887776	11887800	ENSDARG00000069420	11756201	11806249
10	18091276	18091300	ENSDARG00000069377	18017268	18018999
10	18091301	18091325	ENSDARG00000069377	18017268	18018999
10	18091326	18091350	ENSDARG00000069377	18017268	18018999
10	18091351	18091375	ENSDARG00000069377	18017268	18018999
10	18372476	18372500	ENSDARG00000078839	18322191	18497766
10	18372501	18372525	ENSDARG00000078839	18322191	18497766
10	23247276	23247300	ENSDARG00000069333	23218300	23229318
10	23247301	23247325	ENSDARG00000069333	23218300	23229318
10	23247326	23247350	ENSDARG00000069333	23218300	23229318
10	23247351	23247375	ENSDARG00000069333	23218300	23229318
10	23247376	23247400	ENSDARG00000069333	23218300	23229318
10	24740776	24740800	ENSDARG00000078618	24735730	24750722
10	24740801	24740825	ENSDARG00000078618	24735730	24750722
10	24740826	24740850	ENSDARG00000078618	24735730	24750722
10	24740851	24740875	ENSDARG00000078618	24735730	24750722
10	25036426	25036450	ENSDARG00000062347	24995231	25052835
10	25036451	25036475	ENSDARG00000062347	24995231	25052835
10	25036476	25036500	ENSDARG00000062347	24995231	25052835
10	26077526	26077550	ENSDARG00000069139	26037453	26205077
10	26077551	26077575	ENSDARG00000069139	26037453	26205077
10	42818276	42818300	ENSDARG00000033411	42773821	42814933
11	10302301	10302325	ENSDARG00000087133	10036396	10037914
11	10302326	10302350	ENSDARG00000087133	10036396	10037914
11	10302351	10302375	ENSDARG00000087133	10036396	10037914
11	10302376	10302400	ENSDARG00000087133	10036396	10037914
11	10342626	10342650	ENSDARG00000018971	10605252	10614375
11	12382926	12382950	ENSDARG00000076643	12368353	12412556
11	12561976	12562000	ENSDARG00000090567	12569037	12569331
11	12562001	12562025	ENSDARG00000090567	12569037	12569331

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
11	12702451	12702475	ENSDARG00000087708	12705256	12707055
11	12710551	12710575	ENSDARG00000090238	12717744	12718038
11	12710576	12710600	ENSDARG00000090238	12717744	12718038
11	12734101	12734125	ENSDARG00000089182	12734772	12736670
11	12734126	12734150	ENSDARG00000089182	12734772	12736670
11	15034376	15034400	ENSDARG00000019941	15013810	15019264
11	15290351	15290375	ENSDARG00000016405	15300991	15347177
11	15804876	15804900	ENSDARG00000069481	15707808	15797645
11	16140276	16140300	ENSDARG00000040238	16231470	16300511
11	17788376	17788400	ENSDARG00000088837	17563218	17993713
11	17788401	17788425	ENSDARG00000088837	17563218	17993713
11	17788426	17788450	ENSDARG00000088837	17563218	17993713
11	30550926	30550950	ENSDARG00000091126	30550587	30554461
11	30550951	30550975	ENSDARG00000091126	30550587	30554461
11	31938176	31938200	ENSDARG00000086881	31914386	31914898
11	31938201	31938225	ENSDARG00000086881	31914386	31914898
11	31943151	31943175	ENSDARG00000086881	31914386	31914898
11	31968676	31968700	ENSDARG00000075030	32015333	32027918
11	31968701	31968725	ENSDARG00000075030	32015333	32027918
11	32007376	32007400	ENSDARG00000075030	32015333	32027918
11	32007401	32007425	ENSDARG00000075030	32015333	32027918
11	32007851	32007875	ENSDARG00000075030	32015333	32027918
11	39186201	39186225	ENSDARG00000070214	39115699	39161277
12	3057726	3057750	ENSDARG00000089561	2994655	3062179
12	3057751	3057775	ENSDARG00000089561	2994655	3062179
12	3057776	3057800	ENSDARG00000089561	2994655	3062179
12	5891626	5891650	ENSDARG00000086142	5811858	5886480
12	15996251	15996275	ENSDARG00000040712	16028129	16075734
12	15996276	15996300	ENSDARG00000040712	16028129	16075734
12	15996301	15996325	ENSDARG00000040712	16028129	16075734
12	15996326	15996350	ENSDARG00000040712	16028129	16075734
12	15999601	15999625	ENSDARG00000040712	16028129	16075734
12	15999626	15999650	ENSDARG00000040712	16028129	16075734
12	15999651	15999675	ENSDARG00000040712	16028129	16075734
12	15999676	15999700	ENSDARG00000040712	16028129	16075734
12	16476501	16476525	ENSDARG00000079572	16469349	16510140

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
12	24276451	24276475	ENSDARG00000055291	24270740	24282787
12	24276476	24276500	ENSDARG00000055291	24270740	24282787
12	30373401	30373425	ENSDARG00000045092	30373955	30383307
12	30373426	30373450	ENSDARG00000045092	30373955	30383307
12	32013826	32013850	ENSDARG00000091768	32077011	32079801
12	32014751	32014775	ENSDARG00000091768	32077011	32079801
12	37072301	37072325	ENSDARG00000091534	37065405	37083525
12	39478301	39478325	ENSDARG00000086420	39477405	39478529
12	39478326	39478350	ENSDARG00000086420	39477405	39478529
12	50073251	50073275	ENSDARG00000091739	50004532	50075404
12	50073276	50073300	ENSDARG00000091739	50004532	50075404
13	12108176	12108200	ENSDARG00000076127	12169585	12244750
13	12108201	12108225	ENSDARG00000076127	12169585	12244750
13	12108226	12108250	ENSDARG00000076127	12169585	12244750
13	12108251	12108275	ENSDARG00000076127	12169585	12244750
13	14331101	14331125	ENSDARG00000056651	14052306	14348537
13	14331126	14331150	ENSDARG00000056651	14052306	14348537
13	14331151	14331175	ENSDARG00000056651	14052306	14348537
13	16534401	16534425	ENSDARG00000074059	16391899	16546006
13	16534426	16534450	ENSDARG00000074059	16391899	16546006
13	16534451	16534475	ENSDARG00000074059	16391899	16546006
13	19361026	19361050	ENSDARG00000076815	19328842	19362475
13	26811451	26811475	ENSDARG00000093327	26450300	26927657
13	26811476	26811500	ENSDARG00000093327	26450300	26927657
13	26811501	26811525	ENSDARG00000093327	26450300	26927657
13	43005476	43005500	ENSDARG00000009160	42925502	42959184
13	43005501	43005525	ENSDARG00000009160	42925502	42959184
13	43605326	43605350	ENSDARG00000094454	43415207	43664837
13	43605351	43605375	ENSDARG00000094454	43415207	43664837
14	7122676	7122700	ENSDARG00000078572	7109593	7129270
14	7122701	7122725	ENSDARG00000078572	7109593	7129270
14	11543101	11543125	ENSDARG00000086548	11529151	11590125
14	11543126	11543150	ENSDARG00000086548	11529151	11590125
14	11728126	11728150	ENSDARG00000088612	11708991	11727821
14	11728151	11728175	ENSDARG00000088612	11708991	11727821
14	15733876	15733900	ENSDARG00000076332	15444903	15454441

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
14	15733901	15733925	ENSDARG00000076332	15444903	15454441
14	16127801	16127825	ENSDARG00000095624	16124348	16128480
14	16127826	16127850	ENSDARG00000095624	16124348	16128480
14	16913826	16913850	ENSDARG00000078583	16988626	16992583
14	20398651	20398675	ENSDARG00000057494	20056685	20080757
14	20398676	20398700	ENSDARG00000057494	20056685	20080757
14	20398701	20398725	ENSDARG00000057494	20056685	20080757
14	20398726	20398750	ENSDARG00000057494	20056685	20080757
14	24021526	24021550	ENSDARG00000078942	24190167	24481897
14	24021551	24021575	ENSDARG00000078942	24190167	24481897
14	24832251	24832275	ENSDARG00000086954	24833336	24862179
14	25505326	25505350	ENSDARG00000034268	25491482	25760825
14	25505351	25505375	ENSDARG00000034268	25491482	25760825
14	26645726	26645750	ENSDARG00000021250	26600156	26608555
14	30892301	30892325	ENSDARG00000061603	30676730	30775591
14	30892326	30892350	ENSDARG00000061603	30676730	30775591
14	30892351	30892375	ENSDARG00000061603	30676730	30775591
14	46760401	46760425	ENSDARG00000074535	46766598	46812812
15	246251	246275	ENSDARG00000063651	206750	291711
15	7144726	7144750	ENSDARG00000018073	6876279	6893879
15	7190126	7190150	ENSDARG00000018073	6876279	6893879
15	7194126	7194150	ENSDARG00000018073	6876279	6893879
15	7194151	7194175	ENSDARG00000018073	6876279	6893879
15	7198401	7198425	ENSDARG00000018073	6876279	6893879
15	7486776	7486800	ENSDARG00000074969	7350536	7579237
15	11175476	11175500	ENSDARG00000090700	11118222	11129393
15	11189101	11189125	ENSDARG00000090700	11118222	11129393
15	11189126	11189150	ENSDARG00000090700	11118222	11129393
15	11192926	11192950	ENSDARG00000090700	11118222	11129393
15	11256976	11257000	ENSDARG00000090700	11118222	11129393
15	11257326	11257350	ENSDARG00000090700	11118222	11129393
15	16056926	16056950	ENSDARG00000042332	16049488	16067837
15	16056951	16056975	ENSDARG00000042332	16049488	16067837
15	16817476	16817500	ENSDARG00000058263	16838036	16845832
15	17231226	17231250	ENSDARG00000074526	17203687	17302903
15	18032176	18032200	ENSDARG00000075334	17936831	18029159

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
15	18751376	18751400	ENSDARG00000063372	18745015	18755003
15	18751401	18751425	ENSDARG00000063372	18745015	18755003
15	19680501	19680525	ENSDARG00000073872	19614328	19692939
15	21542526	21542550	ENSDARG00000062614	21632389	21662344
15	21542551	21542575	ENSDARG00000062614	21632389	21662344
15	21542576	21542600	ENSDARG00000062614	21632389	21662344
15	21609076	21609100	ENSDARG00000062614	21632389	21662344
15	21848776	21848800	ENSDARG00000062627	21844509	21893734
15	21848801	21848825	ENSDARG00000062627	21844509	21893734
15	25642801	25642825	ENSDARG00000082878	25633436	25633554
15	28998601	28998625	ENSDARG00000068965	29011495	29062453
15	36756176	36756200	ENSDARG00000091388	36770291	36771852
15	36756201	36756225	ENSDARG00000091388	36770291	36771852
15	38124251	38124275	ENSDARG00000078366	38046742	38213653
15	40246976	40247000	ENSDARG00000076093	40242622	40244426
15	45661526	45661550	ENSDARG00000017217	45784978	45942714
15	45661551	45661575	ENSDARG00000017217	45784978	45942714
16	4095651	4095675	ENSDARG00000045959	4049463	4072378
16	21893876	21893900	ENSDARG00000009023	22143616	22239485
16	21899501	21899525	ENSDARG00000009023	22143616	22239485
16	21899526	21899550	ENSDARG00000009023	22143616	22239485
16	23359776	23359800	ENSDARG00000040482	23326921	23503085
16	26516376	26516400	ENSDARG00000040291	26452371	26757333
16	28082726	28082750	ENSDARG00000014975	28079385	28080600
16	29048326	29048350	ENSDARG00000005112	29024966	29053582
16	29048351	29048375	ENSDARG00000005112	29024966	29053582
16	29048451	29048475	ENSDARG00000005112	29024966	29053582
16	29532951	29532975	ENSDARG00000055854	29573622	29597577
16	29532976	29533000	ENSDARG00000055854	29573622	29597577
16	29533001	29533025	ENSDARG00000055854	29573622	29597577
16	30964976	30965000	ENSDARG00000088693	30955622	30982310
16	34552976	34553000	ENSDARG00000003576	34552490	34590500
16	35938226	35938250	ENSDARG00000034643	35969551	36050492
16	35938251	35938275	ENSDARG00000034643	35969551	36050492
16	35938276	35938300	ENSDARG00000034643	35969551	36050492
16	35938301	35938325	ENSDARG00000034643	35969551	36050492

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
16	40144026	40144050	ENSDARG00000080289	40222915	40223032
16	41191551	41191575	ENSDARG00000079570	41119538	41218889
16	45716426	45716450	ENSDARG00000070000	45707109	45710445
16	46286901	46286925	ENSDARG00000076590	46270515	46287081
16	46286926	46286950	ENSDARG00000076590	46270515	46287081
16	46286951	46286975	ENSDARG00000076590	46270515	46287081
16	46286976	46287000	ENSDARG00000076590	46270515	46287081
16	46881576	46881600	ENSDARG00000040123	46671678	46980291
16	46881601	46881625	ENSDARG00000040123	46671678	46980291
16	46881626	46881650	ENSDARG00000040123	46671678	46980291
16	58523401	58523425	ENSDARG00000059760	58473175	58495396
16	58523426	58523450	ENSDARG00000059760	58473175	58495396
17	3968851	3968875	ENSDARG00000024160	3824188	3857276
17	4311701	4311725	ENSDARG00000043799	4346595	4619797
17	4311726	4311750	ENSDARG00000043799	4346595	4619797
17	4311751	4311775	ENSDARG00000043799	4346595	4619797
17	4311776	4311800	ENSDARG00000043799	4346595	4619797
17	10193651	10193675	ENSDARG00000071262	10170226	10206740
17	10193676	10193700	ENSDARG00000071262	10170226	10206740
17	20855501	20855525	ENSDARG00000061682	20816372	20879078
17	22294751	22294775	ENSDARG00000043406	22141428	22331985
17	22294776	22294800	ENSDARG00000043406	22141428	22331985
17	22294801	22294825	ENSDARG00000043406	22141428	22331985
17	22294826	22294850	ENSDARG00000043406	22141428	22331985
17	23730101	23730125	ENSDARG00000088699	23651267	23724511
17	23730151	23730175	ENSDARG00000088699	23651267	23724511
17	23735051	23735075	ENSDARG00000088699	23651267	23724511
17	23735076	23735100	ENSDARG00000088699	23651267	23724511
17	27077976	27078000	ENSDARG00000021896	27033011	27077487
17	27078001	27078025	ENSDARG00000021896	27033011	27077487
17	27078026	27078050	ENSDARG00000021896	27033011	27077487
17	27078051	27078075	ENSDARG00000021896	27033011	27077487
17	31466126	31466150	ENSDARG00000042861	31459694	31542579
17	31466151	31466175	ENSDARG00000042861	31459694	31542579
17	31466176	31466200	ENSDARG00000042861	31459694	31542579
17	32630626	32630650	ENSDARG00000061391	32594439	32632150

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
17	32630651	32630675	ENSDARG00000061391	32594439	32632150
17	34700776	34700800	ENSDARG00000083919	34475837	34475951
17	35420451	35420475	ENSDARG00000043213	35414627	35445893
17	35420476	35420500	ENSDARG00000043213	35414627	35445893
17	35420501	35420525	ENSDARG00000043213	35414627	35445893
17	35420526	35420550	ENSDARG00000043213	35414627	35445893
17	36723876	36723900	ENSDARG00000013266	36684215	36696534
17	37277976	37278000	ENSDARG00000013020	37237614	37324038
17	37278051	37278075	ENSDARG00000013020	37237614	37324038
17	38918576	38918600	ENSDARG00000005179	38908138	38946419
17	38918601	38918625	ENSDARG00000005179	38908138	38946419
17	38918626	38918650	ENSDARG00000005179	38908138	38946419
17	39935351	39935375	ENSDARG00000085506	39939006	39939103
17	39935376	39935400	ENSDARG00000085506	39939006	39939103
17	39936301	39936325	ENSDARG00000085506	39939006	39939103
17	39936326	39936350	ENSDARG00000085506	39939006	39939103
17	39936376	39936400	ENSDARG00000085506	39939006	39939103
17	39936451	39936475	ENSDARG00000085506	39939006	39939103
17	39937201	39937225	ENSDARG00000085506	39939006	39939103
17	39937251	39937275	ENSDARG00000085506	39939006	39939103
17	39941451	39941475	ENSDARG00000070627	39938450	39939604
17	39942251	39942275	ENSDARG00000070627	39938450	39939604
17	39942476	39942500	ENSDARG00000070627	39938450	39939604
17	39942501	39942525	ENSDARG00000070627	39938450	39939604
17	39943451	39943475	ENSDARG00000084920	39946165	39946262
17	39943476	39943500	ENSDARG00000084920	39946165	39946262
17	39943526	39943550	ENSDARG00000084920	39946165	39946262
17	39943551	39943575	ENSDARG00000084920	39946165	39946262
17	39944376	39944400	ENSDARG00000092339	39944059	39946763
17	39944401	39944425	ENSDARG00000092339	39944059	39946763
17	39950026	39950050	ENSDARG00000092339	39944059	39946763
17	39950051	39950075	ENSDARG00000085889	39953333	39953430
17	39951576	39951600	ENSDARG00000085889	39953333	39953430
17	42691576	42691600	ENSDARG00000089717	42668649	42698188
17	42691601	42691625	ENSDARG00000089717	42668649	42698188
17	43296726	43296750	ENSDARG00000053517	43085339	43264853

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
17	43296751	43296775	ENSDARG00000053517	43085339	43264853
17	44756701	44756725	ENSDARG00000076767	44914119	44916049
17	47081326	47081350	ENSDARG00000089502	47079736	47084730
17	47137576	47137600	ENSDARG00000086941	47112972	47114311
17	47137601	47137625	ENSDARG00000086941	47112972	47114311
17	53841576	53841600	ENSDARG00000013047	53799537	53813413
18	28151	28175	ENSDARG00000089832	34108	35439
18	28176	28200	ENSDARG00000089832	34108	35439
18	29901	29925	ENSDARG00000089832	34108	35439
18	3248401	3248425	ENSDARG00000085047	3304392	3304506
18	7188876	7188900	ENSDARG00000063332	7127177	7294650
18	7188901	7188925	ENSDARG00000063332	7127177	7294650
18	7188926	7188950	ENSDARG00000063332	7127177	7294650
18	8148751	8148775	ENSDARG00000042249	7959828	8119764
18	8148776	8148800	ENSDARG00000042249	7959828	8119764
18	13500326	13500350	ENSDARG00000014215	13281050	13621750
18	13500351	13500375	ENSDARG00000014215	13281050	13621750
18	22014176	22014200	ENSDARG00000062178	21955341	22064109
18	23811476	23811500	ENSDARG00000040926	23652345	23806708
18	23811501	23811525	ENSDARG00000040926	23652345	23806708
18	23811526	23811550	ENSDARG00000040926	23652345	23806708
18	24526376	24526400	ENSDARG00000012248	24703309	24715755
18	24526401	24526425	ENSDARG00000012248	24703309	24715755
18	24526426	24526450	ENSDARG00000012248	24703309	24715755
18	27295626	27295650	ENSDARG00000017173	27282361	27303562
18	27532226	27532250	ENSDARG00000026070	27520356	27575299
18	27532251	27532275	ENSDARG00000026070	27520356	27575299
18	27532276	27532300	ENSDARG00000026070	27520356	27575299
18	33838326	33838350	ENSDARG00000023797	33835785	34021041
18	33838351	33838375	ENSDARG00000023797	33835785	34021041
18	34944526	34944550	ENSDARG00000061742	34941392	34971130
18	41694276	41694300	ENSDARG00000086034	41670067	41714861
18	41694301	41694325	ENSDARG00000086034	41670067	41714861
18	41694326	41694350	ENSDARG00000086034	41670067	41714861
18	41694351	41694375	ENSDARG00000086034	41670067	41714861
19	8313201	8313225	ENSDARG00000036767	8304645	8313511

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
19	17417701	17417725	ENSDARG00000035994	17388307	17483418
19	20782551	20782575	ENSDARG00000062688	20730991	20752475
19	20782576	20782600	ENSDARG00000062688	20730991	20752475
19	20782601	20782625	ENSDARG00000062688	20730991	20752475
19	27347101	27347125	ENSDARG00000062208	27337706	27354388
19	27347126	27347150	ENSDARG00000062208	27337706	27354388
19	27347151	27347175	ENSDARG00000062208	27337706	27354388
19	27347176	27347200	ENSDARG00000062208	27337706	27354388
19	31819976	31820000	ENSDARG00000070491	31767679	31770350
19	31820001	31820025	ENSDARG00000070491	31767679	31770350
19	31820026	31820050	ENSDARG00000070491	31767679	31770350
19	37683351	37683375	ENSDARG00000001559	37663206	38223383
19	37683376	37683400	ENSDARG00000001559	37663206	38223383
19	37683401	37683425	ENSDARG00000001559	37663206	38223383
19	37683426	37683450	ENSDARG00000001559	37663206	38223383
19	38651026	38651050	ENSDARG00000070362	38667444	38668787
19	42535476	42535500	ENSDARG00000078557	42523474	42563363
19	42535501	42535525	ENSDARG00000078557	42523474	42563363
19	42535526	42535550	ENSDARG00000078557	42523474	42563363
19	42535551	42535575	ENSDARG00000078557	42523474	42563363
20	1945226	1945250	ENSDARG00000089996	1985719	2003922
20	15984676	15984700	ENSDARG00000026519	15973492	16196883
20	15984701	15984725	ENSDARG00000026519	15973492	16196883
20	15984726	15984750	ENSDARG00000026519	15973492	16196883
20	15984751	15984775	ENSDARG00000026519	15973492	16196883
20	17475801	17475825	ENSDARG00000085736	17467472	17467586
20	18637551	18637575	ENSDARG00000095347	18623817	18667719
20	18637576	18637600	ENSDARG00000095347	18623817	18667719
20	20698476	20698500	ENSDARG00000021143	20710026	20768113
20	24821076	24821100	ENSDARG00000095716	25121867	25124639
20	24908201	24908225	ENSDARG00000004635	24847468	24992645
20	24908226	24908250	ENSDARG00000004635	24847468	24992645
20	24908251	24908275	ENSDARG00000004635	24847468	24992645
20	24908276	24908300	ENSDARG00000004635	24847468	24992645
20	26619551	26619575	ENSDARG00000055610	26602063	26653649
20	26619576	26619600	ENSDARG00000055610	26602063	26653649

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
20	34130326	34130350	ENSDARG00000017634	34108939	34115348
20	34130351	34130375	ENSDARG00000017634	34108939	34115348
20	40906151	40906175	ENSDARG00000041736	40793454	40798083
20	51990076	51990100	ENSDARG00000091595	51986863	51993973
20	54874776	54874800	ENSDARG00000068912	54848858	54888082
20	54874801	54874825	ENSDARG00000068912	54848858	54888082
20	55778426	55778450	ENSDARG00000087890	55778118	55778271
21	5410476	5410500	ENSDARG00000092263	5404334	5431995
21	5410501	5410525	ENSDARG00000092263	5404334	5431995
21	11753301	11753325	ENSDARG00000060025	11657405	11737698
21	11753326	11753350	ENSDARG00000060025	11657405	11737698
21	14387326	14387350	ENSDARG00000004451	14375999	14422468
21	14387351	14387375	ENSDARG00000004451	14375999	14422468
21	14387376	14387400	ENSDARG00000004451	14375999	14422468
21	14387401	14387425	ENSDARG00000004451	14375999	14422468
21	16153126	16153150	ENSDARG00000011459	16084006	16114767
21	17994951	17994975	ENSDARG00000085209	18091281	18091397
21	17994976	17995000	ENSDARG00000085209	18091281	18091397
21	17995001	17995025	ENSDARG00000085209	18091281	18091397
21	17995026	17995050	ENSDARG00000085209	18091281	18091397
21	21827251	21827275	ENSDARG00000091888	21821235	21824233
21	21827276	21827300	ENSDARG00000091888	21821235	21824233
21	21827301	21827325	ENSDARG00000044612	21827289	21830279
21	21827326	21827350	ENSDARG00000044612	21827289	21830279
21	30551501	30551525	ENSDARG00000011171	30605204	30774786
21	30551526	30551550	ENSDARG00000011171	30605204	30774786
21	30551551	30551575	ENSDARG00000011171	30605204	30774786
21	30551576	30551600	ENSDARG00000011171	30605204	30774786
21	32307426	32307450	ENSDARG00000089914	32470757	32515785
21	32642026	32642050	ENSDARG00000091943	32724279	32733414
21	32642051	32642075	ENSDARG00000091943	32724279	32733414
22	1213426	1213450	ENSDARG00000009933	1189220	1220491
22	4489126	4489150	ENSDARG00000092732	4461006	4463734
22	9222401	9222425	ENSDARG00000093761	9225731	9254366
22	9222426	9222450	ENSDARG00000093761	9225731	9254366
22	9257026	9257050	ENSDARG00000079986	9299141	9322726

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
22	9257051	9257075	ENSDARG00000079986	9299141	9322726
22	11514301	11514325	ENSDARG00000032087	11630399	11642910
22	15359026	15359050	ENSDARG00000094089	15455711	15457528
22	15359051	15359075	ENSDARG00000094089	15455711	15457528
22	18102026	18102050	ENSDARG00000005783	18088959	18246509
22	18102051	18102075	ENSDARG00000005783	18088959	18246509
22	26881326	26881350	ENSDARG00000041538	26881458	26882104
22	26881351	26881375	ENSDARG00000041538	26881458	26882104
22	26881376	26881400	ENSDARG00000041538	26881458	26882104
22	27004676	27004700	ENSDARG00000090014	27058713	27091556
22	27004701	27004725	ENSDARG00000090014	27058713	27091556
22	27004726	27004750	ENSDARG00000090014	27058713	27091556
22	31440026	31440050	ENSDARG00000071058	31445545	31446853
22	31442101	31442125	ENSDARG00000071058	31445545	31446853
22	31442126	31442150	ENSDARG00000071058	31445545	31446853
22	31509826	31509850	ENSDARG00000071059	31414989	31428551
22	33494851	33494875	ENSDARG00000040920	33447236	33569743
22	34550651	34550675	ENSDARG00000089663	34461570	34472027
22	34550676	34550700	ENSDARG00000089663	34461570	34472027
22	35562751	35562775	ENSDARG00000093991	35545918	35547105
23	4823501	4823525	ENSDARG00000045945	4807675	4831162
23	17014526	17014550	ENSDARG00000037645	17001781	17020233
23	17014551	17014575	ENSDARG00000037645	17001781	17020233
23	19953801	19953825	ENSDARG00000070894	19943851	19956308
23	23397276	23397300	ENSDARG00000077852	23314851	23418043
23	23397301	23397325	ENSDARG00000077852	23314851	23418043
23	23397326	23397350	ENSDARG00000077852	23314851	23418043
23	25373701	25373725	ENSDARG00000007436	25326334	25337860
23	32671201	32671225	ENSDARG00000036826	32600664	32629321
23	32671226	32671250	ENSDARG00000036826	32600664	32629321
23	32671251	32671275	ENSDARG00000036826	32600664	32629321
23	32671276	32671300	ENSDARG00000036826	32600664	32629321
23	34780026	34780050	ENSDARG00000077437	34615326	34769564
23	38008201	38008225	ENSDARG00000091015	37935481	38019298
23	38008226	38008250	ENSDARG00000091015	37935481	38019298
23	38008251	38008275	ENSDARG00000091015	37935481	38019298

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
23	39398076	39398100	ENSDARG00000090234	39375812	39384281
23	39398551	39398575	ENSDARG00000090234	39375812	39384281
23	39785276	39785300	ENSDARG00000075468	39778184	39812043
23	41437026	41437050	ENSDARG00000070227	41297168	41506870
24	8286851	8286875	ENSDARG00000071774	8139814	8157025
24	8286876	8286900	ENSDARG00000071774	8139814	8157025
24	10571126	10571150	ENSDARG00000084053	10618690	10618815
24	10571151	10571175	ENSDARG00000084053	10618690	10618815
24	10571176	10571200	ENSDARG00000084053	10618690	10618815
24	10571201	10571225	ENSDARG00000084053	10618690	10618815
24	16081176	16081200	ENSDARG00000096455	16078332	16081433
24	17017926	17017950	ENSDARG00000058821	17020682	17256043
24	17017951	17017975	ENSDARG00000058821	17020682	17256043
24	20011176	20011200	ENSDARG00000038428	20067997	20136956
24	20011201	20011225	ENSDARG00000038428	20067997	20136956
24	23158876	23158900	ENSDARG00000062415	22730185	23258168
24	27030201	27030225	ENSDARG00000062228	27037668	27075454
24	30523676	30523700	ENSDARG00000088190	30522348	30524684
24	33401901	33401925	ENSDARG00000034518	33357085	33401868
24	33402901	33402925	ENSDARG00000034518	33357085	33401868
24	37827126	37827150	ENSDARG00000025081	37836254	37861603
24	38141126	38141150	ENSDARG00000075519	38080714	38110063
24	38561451	38561475	ENSDARG00000091372	38522346	38593922
24	42664526	42664550	ENSDARG00000087859	42646371	42659643
25	3699351	3699375	ENSDARG00000060674	3698884	3703316
25	3699376	3699400	ENSDARG00000060674	3698884	3703316
25	3699401	3699425	ENSDARG00000060674	3698884	3703316
25	12685526	12685550	ENSDARG00000096466	12624606	12629657
25	13463651	13463675	ENSDARG00000090312	13433477	13434067
25	17060326	17060350	ENSDARG00000070717	16973622	16996898
25	19064151	19064175	ENSDARG00000052004	19050766	19180696
25	19064176	19064200	ENSDARG00000052004	19050766	19180696
25	19064201	19064225	ENSDARG00000052004	19050766	19180696
25	19064226	19064250	ENSDARG00000052004	19050766	19180696
25	19319376	19319400	ENSDARG00000062960	19368324	19378456
25	19319401	19319425	ENSDARG00000062960	19368324	19378456

Chr	Region Start	Region End	Nearest Gene	Gene Start	Gene End
25	19319426	19319450	ENSDARG00000062960	19368324	19378456
25	19354001	19354025	ENSDARG00000062960	19368324	19378456
25	20325476	20325500	ENSDARG00000058259	20244191	20356971
25	24511326	24511350	ENSDARG00000045696	24500808	24539288
25	25944826	25944850	ENSDARG00000040854	25864839	25868121
25	25944851	25944875	ENSDARG00000040854	25864839	25868121
25	25944876	25944900	ENSDARG00000040854	25864839	25868121
25	25984576	25984600	ENSDARG00000040854	25864839	25868121
25	25984601	25984625	ENSDARG00000040854	25864839	25868121
25	26001326	26001350	ENSDARG00000040854	25864839	25868121
25	26004951	26004975	ENSDARG00000040854	25864839	25868121
25	26004976	26005000	ENSDARG00000040854	25864839	25868121
25	27098501	27098525	ENSDARG00000096101	27097208	27097595
25	29111876	29111900	ENSDARG00000013312	28804289	29136657
25	30224151	30224175	ENSDARG00000074419	30195791	30296040
25	30848226	30848250	ENSDARG00000029431	31083584	31103883
25	30848251	30848275	ENSDARG00000029431	31083584	31103883
25	30848276	30848300	ENSDARG00000029431	31083584	31103883